TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Dani Korpi

Full-Duplex Wireless: Self-interference Modeling, Digital Cancellation, and System Studies

Julkaisu 1516 • Publication 1516

Tampere 2017

Tampereen teknillinen yliopisto. Julkaisu 1516
Tampere University of Technology. Publication 1516

Dani Korpi

Full-Duplex Wireless: Self-interference Modeling, Digital Cancellation, and System Studies

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 1st of December 2017, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2017

Doctoral candidate: Dani Korpi
Laboratory of Electronics and Communications Engineering
Faculty of Computing and Electrical Engineering
Tampere University of Technology
Tampere, Finland

Supervisor: Mikko Valkama, Professor
Laboratory of Electronics and Communications Engineering
Faculty of Computing and Electrical Engineering
Tampere University of Technology
Tampere, Finland

Instructors: Taneli Riihonen, Assistant Professor
Laboratory of Electronics and Communications Engineering
Faculty of Computing and Electrical Engineering
Tampere University of Technology
Tampere, Finland

Lauri Anttila, D.Sc. (Tech.)
Laboratory of Electronics and Communications Engineering
Faculty of Computing and Electrical Engineering
Tampere University of Technology
Tampere, Finland

Pre-examiners: Melissa Duarte, Ph.D.
Mathematical and Algorithmic Sciences Lab
France Research Center, Huawei Technologies
Paris, France

Himal A. Suraweera, Senior Lecturer
Department of Electrical and Electronic Engineering
University of Peradeniya
Peradeniya, Sri Lanka

Opponent: Daniel W. Bliss, Associate Professor
School of Electrical, Computer, and Energy Engineering
Arizona State University
Tempe, Arizona, USA

ISBN 978-952-15-4052-3 (printed)
ISBN 978-952-15-4059-2 (PDF)
ISSN 1459-2045

Abstract

In the recent years, a significant portion of the research within the field of wireless communications has been motivated by two aspects: the constant increase in the number of wireless devices and the higher and higher data rate requirements of the individual applications. The undisputed outcome of these phenomena is the heavy congestion of the suitable spectral resources. This has inspired many innovative solutions for improving the spectral efficiency of the wireless communications systems by facilitating more simultaneous connections and higher data rates without requiring additional spectrum. These include technologies such as in-phase/quadrature (I/Q) modulation, multiple-input and multiple-output (MIMO) systems, and the orthogonal frequency-division multiplexing (OFDM) waveform, among others. Even though these existing solutions have greatly improved the spectral efficiency of wireless communications, even more advanced techniques are needed for fulfilling the future data transfer requirements in the ultra high frequency (UHF) band, which is perhaps the most congested piece of the spectrum. To this end, wireless inband full-duplex (IBFD) communications has recently been proposed as another step towards the full utilization of the spectral resources. What IBFD simply refers to is simultaneous transmission and reception on the same frequency band within the same device, meaning that the spectral efficiency is in theory doubled. Considering the cost of the radio frequency bands from a teleoperator's perspective, IBFD can therefore provide savings in the order of hundreds of millions while also easing the congestion of the UHF and lower bands. Hence, implementing and employing radio transceivers capable of IBFD communications is a highly tempting vision.

Firstly, this thesis concentrates on solving the greatest challenge in wireless IBFD communications: the self-interference (SI). In particular, SI refers to the interference produced by the own transmitter, which is directly coupling to the receiver. Since the transmitted and received signals within an IBFD transceiver are fully overlapping in the frequency domain, conventional duplex filtering is of no use in suppressing the SI, and hence advanced techniques are needed for canceling it. The required SI cancellation solutions can therefore be considered the cost of the improved spectral efficiency. To this end, this thesis first identifies which of the radio circuit impairments must be considered when modeling and canceling the SI within an IBFD transceiver. Especially,

the nonlinear distortion produced by the transmitter power amplifier is shown to be one of the dominant impairments. These findings are then used as a basis for developing SI cancellation algorithms operating in the *digital domain* of the receiver. To achieve sufficiently high modeling accuracy, the proposed digital SI cancellers take into account the effects of the analog impairments that distort the SI signal while it is propagating to the receiver. As a result, the SI signal models derived in this thesis are some of the most comprehensive reported in the literature. In addition, two alternative schemes are presented for efficiently estimating the necessary parameters needed for regenerating the SI signal.

The proposed digital cancellation algorithms are then evaluated by utilizing them in an actual IBFD prototype, which contains also other SI suppression mechanisms operating in the analog domain. The obtained measurement results show that, with the help of these digital SI cancellers, the SI can be cancelled almost perfectly, proving that true IBFD operation is indeed possible. In particular, having canceled the SI completely, the receiver will experience no additional interference produced by the simultaneously operating transmitter, and hence the overall data rate can be doubled without introducing any additional spectral resources. Altogether, the own transmit signal is shown to be suppressed in some cases by more than 100 dB, which is one of the highest reported SI cancellation performances to date.

Secondly, this thesis also provides a comprehensive system level analysis of a network with an IBFD-capable access node. Especially, it is assumed that the access node serves legacy half-duplex downlink and uplink mobile users simultaneously on the same frequency band while also using the same spectral resources for backhauling all the data wirelessly. The system is analyzed in terms of solving the optimal transmit power allocation under minimum data rate requirements. By comparing the IBFD access node to two reference scenarios where either all or some of the communication tasks are divided in time, it is shown that the IBFD capability is highly beneficial in many respects. This indicates that the gains of the IBFD technology can be realized to some extent even when most of the devices are only half-duplex capable.

Altogether, the findings of this thesis show that wireless IBFD communications is indeed possible in practice, as long as all the significant analog impairments are considered in the digital SI cancellation stage. Moreover, guidelines for implementing IBFD transceivers under practical constraints are provided while also demonstrating that the IBFD technology results in various network-level benefits. These results pave the way towards the commercial deployment of IBFD radio transceivers.

This dissertation is a compound thesis that consists of a summary and seven original journal publications. In addition, it reports research results from one transactions manuscript and 14 conference papers where the author of the thesis is the leading contributor, as well as two original journal publications, one book chapter, and seven conference papers where the author of this thesis is a co-author participating in the research work. However, for brevity, only the seven journal publications are attached to this thesis.

Preface

This thesis is based on the research work carried out during the years 2014–2016 in the Laboratory of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland. I would like to gratefully acknowledge the financial support I received from the Tampere University of Technology Graduate School (during 2016–2017), Nokia Foundation, Tuula and Yrjö Neuvo Research Fund, Emil Aaltonen Foundation, and Pekka Ahonen Fund. In addition, I also wish to acknowledge the funding received from Academy of Finland (under the projects #259915 "In-Band Full-Duplex MIMO Transmission: A Breakthrough to High-Speed Low-Latency Mobile Networks", #301820 "Competitive Funding to Strengthen University Research Profiles", and #304147 "In-Band Full-Duplex Radio Technology: Realizing Next Generation Wireless Transmission"), Finnish Funding Agency for Technology and Innovation (Tekes, under the projects "Full-Duplex Cognitive Radio" and "TAKE-5"), and Intel Corporation.

First and foremost, I wish to thank my supervisor Prof. Mikko Valkama, who has been one of my academic role models already since my time as a research assistant. His hard work and dedication have inspired me to push also my own boundaries in order to take the full advantage of my potential, for which I am forever thankful. I must also thank Prof. Markku Renfors, who instructed and guided me during my first years at the department. I feel truly privileged to have been able to learn from his extensive experience and knowledge within the field of wireless communications. In addition, I have received excellent guidance from D.Sc. Taneli Riihonen, who has spared no effort in teaching me the intricacies of scientific research and writing. I am thankful for having had such an exceptional talent as my instructor. Furthermore, this thesis could not have been made without the help and supervision of D.Sc. Lauri Anttila, whom I also wish to thank. His experience and knowledge regarding the modeling of radio transceivers helped me to set the foundation for the research work reported in this thesis.

I am also very grateful to Dr. Melissa Duarte and Dr. Himal A. Suraweera for acting as the pre-examiners of this thesis. Their invaluable comments and insights helped to improve this thesis even further. In addition, I wish to thank Prof. Daniel W. Bliss for agreeing to act as the opponent in the public examination of this thesis.

I must also acknowledge the various collaboration partners that have been co-operating with me throughout my research work. Firstly, I wish to thank Prof. Risto Wichman, Prof. Katsuyuki Haneda, D.Sc. Clemens Icheln, Mikko Heino, Emilio Antonio-Rodríguez, Sathya Venkatasubramanian and many others from Aalto University for the numerous years of research collaboration and various co-authored publications. Moreover, I would like to thank Dr. Shilpa Talwar, Dr. Yang-Seok Choi, and Timo Huusari from Intel Corporation for the fruitful research collaboration without which this thesis could not have been made. Also, I must thank Prof. Ashutosh Sabharwal and Prof. Joseph Cavallaro from Rice University for giving me an opportunity to visit their research groups in late 2014. I learned a great deal during this research visit, and thoroughly enjoyed my time in Houston, Texas.

In addition, I wish to thank my numerous co-workers in the research group. The efforts of Joose Tamminen and Matias Turunen have been invaluable for my research work, as they have taken care of many practical issues regarding the prototype implementation and measurement setup. Furthermore, I would also like to thank Ville Syrjälä for providing guidance with various research problems. I also wish to acknowledge the almost legendary 10.30 lunch group, and its past and present participants: Janis Werner, Aki Hakkarainen, Joonas Säe, and Mike Koivisto. Without you, the daily lunch break would have been so much more boring. Moreover, there are also various other people with whom I have had countless great discussions regarding work and life in general that I want to thank. They include Jaakko Marttila, Markus Allén, Jukka Talvitie, Mahmoud Abdelaziz, Pedro Silva, Toni Levanen, Adnan Kiayani, Paschalis Sofotasios, Simran Singh, Ahmet Gökceoglu, Muhammad Zeeshan Waheed, Vesa Lehtinen, Sari Kinnari, and Tuija Grek. Each of you has been an integral part in creating the great spirit within our research group.

There are also so many friends whose names I have not mentioned here but whom I also must thank. Especially, I want to mention all the people I have got to know via Tampere Pentecostal Church. With your help, I have always been able to keep my bearings and remember what is truly important in life. I would also like to express my most sincere gratitude to Radiohead, Jean Sibelius, and the National, among many others, for creating the soundtrack of these past years and months. Furthermore, I am also thankful to my nephew and two nieces, Taito, Mai, and Sivia, who have always been able to make me forget my worries by just being who they are.

Lastly, I would like to offer my most sincere thanks and gratitude to my parents Riitta and Väinö, as well as to my siblings Markus, Mika, Säde, and Jimi. I firmly believe that everything I have presented in this thesis builds on the solid foundation that you helped me to set already during my childhood. Without the self-discipline and the humble attitude that I have learned from you, I would not be here writing this today.

Tampere, October 2017
Dani Korpi

Table of Contents

List of Publications

This thesis is a compound thesis based on the following seven publications.

[**P1**] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.

[**P2**] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.

[**P3**] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodríguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full-duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, May 2015.

[**P4**] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: Pushing the limits," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 80–87, Sep. 2016.

[**P5**] D. Korpi, M. Heino, C. Icheln, K. Haneda, and M. Valkama, "Compact inband full-duplex relays with beyond 100 dB self-interference suppression: Enabling techniques and field measurements," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 2, pp. 960–965, Feb. 2017.

[**P6**] D. Korpi, L. Anttila, and M. Valkama, "Nonlinear self-interference cancellation in MIMO full-duplex transceivers under crosstalk," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, Feb. 2017.

[**P7**] D. Korpi, T. Riihonen, A. Sabharwal, and M. Valkama, "Transmit power optimization and feasibility analysis of self-backhauling full-duplex radio access systems," *IEEE Transactions on Wireless Communications*, under review after revision, 2017.

Furthermore, there are also several related works, particularly **[4, 17, 18, 82, 99, 120–134, 160, 193, 194, 233, 238]**[1], that have been published as a result of this dissertation work and research, although the compound thesis format does not permit attaching them to the present document.

[1]The publications which the author of this thesis has (co-)authored are cited using boldface numbers throughout this thesis.

Abbreviations

5G	Fifth generation
ADC	Analog-to-digital converter
ADSL	Asymmetric digital subscriber line
AGC	Automatic gain control
AN	Access node
BN	Backhaul node
BPF	Band-pass filter
BS	Base station
CDF	Cumulative distribution function
CFO	Carrier frequency offset
CPE	Common phase error
CSI	Channel state information
CSM	Command and service module
CSMA/CA	Carrier sense multiple access with collision avoidance
CW	Continuous wave
DAC	Digital-to-analog converter
DC	Direct current
DL	Downlink
DLS	Dithered linear search
DVB	Digital video broadcasting
EBD	Electrical balance duplexer
EVM	Error vector magnitude
FDD	Frequency-division duplex
FIR	Finite impulse response

FPGA	Field-programmable gate array
HD	Half-duplex
I/Q	In-phase/quadrature
IBFD	Inband full-duplex
IIP2	2nd-order input intercept point
IIP3	3rd-order input intercept point
IRR	Image rejection ratio
ITU	International Telecommunication Union
IUI	Inter-user-interference
LMS	Least mean squares
LNA	Low-noise amplifier
LO	Local oscillator
LOS	Line-of-sight
LPF	Low-pass filter
LS	Least squares
LTE	Long term evolution
MAC	Medium access control
MIMO	Multiple-input and multiple-output
MISO	Multiple-input and single-output
MRA	Multi-reconfigurable antenna
NF	Noise figure
NI	National Instruments
NLOS	Non-line-of-sight
OFDM	Orthogonal frequency-division multiplexing
PA	Power amplifier
PAPR	Peak-to-average power ratio
PCA	Principal component analysis
PH	Parallel Hammerstein
PLL	Phase-locked loop
PSD	Power spectral density
QoS	Quality of Service
RF	Radio frequency
RLS	Recursive least squares
RRM	Radio resource management
RX	Receiver
SI	Self-interference
SINR	Signal-to-interference-plus-noise ratio
SISO	Single-input and single-output

SQNR	Signal-to-quantization-noise ratio
STAR	Simultaneous transmission and reception
SVD	Singular value decomposition
TDD	Time-division duplex
TX	Transmitter
UE	User equipment
UHF	Ultra high frequency
UL	Uplink
VGA	Variable-gain amplifier
VM	Vector modulator
VST	Vector signal transceiver
ZF	Zero-forcing

Symbols

$\mathbf{1}_N$	Column vector consisting of N ones
a_{ANT}	Amount of passive SI isolation
a_{DC}	Amount of digital cancellation
a_{RFC}	Amount of active RF cancellation
b	Number of bits in the RX ADC
$B(t)$	Brownian motion process
C_{FD}	Capacity of an IBFD communications system
C_{HD}	Capacity of a HD communications system
$c(n)$	Digital cancellation signal in a SISO transceiver
$c_i(n)$	Digital cancellation signal in the ith receiver
D	Number of DL UEs
e	Euler's number
$e_{i,\mathrm{tot}}(n)$	Modeling error signal before digital cancellation in the ith receiver
$\mathbf{e}_i$	Modeling error vector before digital cancellation in the ith receiver
$e_{i,\mathrm{RX}}(t)$	Modeling error and noise signal in the ith receiver
$e_{i,\mathrm{TX}}$ (t)	Modeling error and noise signal in the jth transmitter
f	Frequency variable
$f_{i,\mathrm{RX}}$	Linear response of the ith receiver
$f_{j,\mathrm{TX}}$	Linear response of the jth transmitter
F_{rx}	Noise factor of the receiver
F_{tx}	Noise factor of the transmitter
$\widehat{\mathbf{h}}_0$	The initial SI channel estimate of the complexity reduction scheme
$\mathbf{h}_i(m)$	Coefficient vector of basis functions with lag m in the ith receiver
$\mathbf{h}_i$	Complete SI channel vector in the ith receiver

$\widehat{\mathbf{h}}_i$	Estimate of the complete SI channel vector in the ith receiver
$\widehat{\mathbf{h}}_i^{\mathrm{LMS}}$	LMS estimate of the complete SI channel vector in the ith receiver
$\widehat{\mathbf{h}}_i^{\mathrm{LS}}$	LS estimate of the complete SI channel vector in the ith receiver
$\hat{h}^{\mathrm{L}}$	One-tap estimate of the linear SI channel
$h_{ij}^{\mathrm{L}}(m)$	Total channel response (linear signal model)
$h_{ij,p}^{\mathrm{NL}}(m)$	Total pth-order channel response (nonlinear signal model)
$h_{i,\boldsymbol{\xi}^k}^{\mathrm{NL}}(m)$	Total channel response of $\psi_{\boldsymbol{\xi}^k}(x_1(n),\ldots,x_{N_{\mathrm{t}}}(n))$
$h_{ij,1}^{\mathrm{WL}}(m)$	Total linear channel response (widely linear signal model)
$h_{ij,2}^{\mathrm{WL}}(m)$	Total image channel response (widely linear signal model)
$h_p(m)$	Coefficients of the pth-order term in the PH model
$h_{ij,\mathrm{RFC}}(t)$	Channel estimate used for RF cancellation in the ith receiver
$h_{ij,\mathrm{RSI}}(t)$	Effective SI channel after RF cancellation in the ith receiver
$h_{\mathrm{SI}}(t)$	SISO coupling channel between the transmitter and the receiver
$h_{ij,\mathrm{SI}}(t)$	Coupling channel between the jth transmitter and the ith receiver
$\mathbf{I}_K$	$K \times K$ identity matrix
$iip2_{\mathrm{MX}}$	2nd-order input intercept point of the RX I/Q mixer
$iip2_{\mathrm{VGA}}$	2nd-order input intercept point of the RX VGA
$iip3_{\mathrm{LNA}}$	3rd-order input intercept point of the LNA
$iip3_{\mathrm{MX}}$	3rd-order input intercept point of the RX I/Q mixer
$iip3_{\mathrm{PA}}$	3rd-order input intercept point of the PA
$iip3_{\mathrm{VGA}}$	3rd-order input intercept point of the RX VGA
$irr_{\mathrm{RX}}(f)$	Frequency-dependent RX IRR
$irr_{\mathrm{TX}}(f)$	Frequency-dependent TX IRR
K	Total number of static basis functions
$k_{1,\mathrm{RX}}(t)$	Response of the linear component in the receiver
$k_{1,\mathrm{TX}}(t)$	Response of the linear component in the transmitter
$k_{2,\mathrm{RX}}(t)$	Response of the image component in the receiver
$k_{2,\mathrm{TX}}(t)$	Response of the image component in the transmitter
$k_{j,p,\mathrm{TX}}(t)$	Response of the pth-order SI term in the jth transmitter
$k_{j,\boldsymbol{\xi}^k,\mathrm{TX}}(t)$	Response of the pth-order MIMO SI term in the jth transmitter
k_{LNA}	Complex voltage gain of the LNA
$k_{i,1,\mathrm{RX}}(t)$	Response of the linear component in the ith receiver
$k_{j,1,\mathrm{TX}}(t)$	Response of the linear component in the jth transmitter
$k_{i,2,\mathrm{RX}}(t)$	Response of the image component in the ith receiver
$k_{j,2,\mathrm{TX}}(t)$	Response of the image component in the jth transmitter
k_{PA}	Complex voltage gain of the PA
k_{VGA}	Complex voltage gain of the VGA
$K_{1,\mathrm{TX}}(f)$	Frequency response of the linear component in the transmitter

$K_{2,\mathrm{TX}}(f)$	Frequency response of the image component in the transmitter
L	Number of basis functions after the PCA-based complexity reduction
L_{B}	Path loss between the BN and the AN
$\mathbf{L}_{\mathrm{Bu}}$	Vector containing the path losses between the BN and the UL UEs
L_i^{Bd}	Path loss between the BN and the ith DL UE
L_j^{Bu}	Path loss between the BN and the jth UL UE
L_i^{d}	Path loss between the AN and the ith DL UE
L_j^{u}	Path loss between the AN and the jth UL UE
$\mathbf{L}_{\mathrm{ud}}$	Matrix containing the path losses between the DL and UL UEs
$\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}$	Ratios of the DL to UL UEs and the DL UEs to AN path losses
L_{ij}^{ud}	Path loss between the ith DL and jth UL UE
M_{PH}	Memory length of the PH model
M	Total number of memory taps
M_1	Number of pre-cursor taps
M_2	Number of post-cursor taps
$M_{\mathrm{r}}^{\mathrm{B}}$	Number of transmitted backhaul data streams
$M_{\mathrm{t}}^{\mathrm{B}}$	Number of received backhaul data streams
$\mathcal{N}\left(\mu, \sigma^2\right)$	Normal distribution with mean μ and variance σ^2
N	Parameter estimation sample size
n	Sample index
N_{r}	Number of RX antennas
N_{t}	Number of TX antennas
N_{RFC}	Number of taps in the RF canceller
$\mathcal{O}(\cdot)$	The Big O operator.
P	Nonlinearity order of the PH model
$\mathbf{p}$	Vector containing all the DL and UL transmit powers
p_{ADC}	Average power in the RX ADC input.
$p_{\mathrm{ADC}}^{\max}$	Peak power in the RX ADC input.
$P_{\mathrm{d}}^{\mathrm{B}}$	Total transmit power allocated for backhauling the DL data
$P_{\mathrm{u}}^{\mathrm{B}}$	Total transmit power allocated for backhauling the UL data
$\mathbf{p}_{\mathrm{d}}$	Vector containing all the DL transmit powers
p_i^{d}	Transmit power allocated for the ith DL signal
$p_{\mathrm{IMD,PA}}$	Power of the PA-induced nonlinear SI term
$p_{\mathrm{IMD,RX}}$	Power of the nonlinearities produced in the RX chain
$p_{\mathrm{n,RX}}$	Power of the RX-induced noise
$p_{\mathrm{n,TX}}$	Power of the TX-induced noise
p_{PN}	Power of the phase-noise-induced SI term
p_{QN}	Power of the quantization noise produced in the RX ADCs

p_{RFC}	Noise power at RF canceller output
p_{RX}	Power of the signal of interest at the receiver input
p_{SI}	Power of the linear SI term
$p_{\mathrm{SI,IM}}$	Power of the SI I/Q image component
p_{SOI}	Power of the signal of interest
p_{th}	Power of the thermal noise for a given bandwidth
p_{TX}	Transmit power
$\mathbf{p}_{\mathrm{u}}$	Vector containing all the UL transmit powers
p_j^{u}	Transmit power of the jth UL UE
p_{x}	Power of TX DAC output signal
$papr$	PAPR of the signal in linear power units
$\mathbf{q}_{\mathrm{B/d}}$	Ratios of the BN to DL UEs and the AN to DL UEs path losses
$\mathbf{q}_{\mathrm{d}}$	Vector containing the inverses of the DL path losses
$\mathbf{q}_{\mathrm{u}}$	Vector containing the inverses of the UL path losses
R	Impedance of the system
R^{d}	DL sum-rate of the system
$R_{\mathrm{B}}^{\mathrm{d}}$	Overall rate for backhauling DL data
R_i^{d}	Data rate of the ith DL UE
R^{u}	UL sum-rate of the system
$R_{\mathrm{B}}^{\mathrm{u}}$	Overall rate for backhauling UL data
R_j^{u}	Data rate of the jth UL UE
$\mathbf{R}_{\boldsymbol{\psi}}$	Covariance matrix of the instantaneous basis function vector $\boldsymbol{\psi}(n)$
$S_{\mathrm{B/d}}$	Sum of the elements of $\mathbf{q}_{\mathrm{B/d}}$
$S_{\mathrm{B/u}}$	Sum defined by $\mathbf{L}_{\mathrm{Bu}}^T\mathbf{q}_{\mathrm{u}}$
S_{d}	Sum of the inverted DL path losses (elements of $\mathbf{q}_{\mathrm{d}}$)
S_{ud}	Sum of the elements of $\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}\mathbf{q}_{\mathrm{u}}$
$\mathbf{S}_{\mathrm{PCA}}$	Complete PCA transformation matrix
$\mathbf{s}_{\mathrm{PCA},n}$	The nth column of $\mathbf{S}_{\mathrm{PCA}}$
$\widetilde{\mathbf{S}}_{\mathrm{PCA}}$	Reduced PCA transformation matrix
$\mathbf{S}_{\boldsymbol{\psi}}$	Orthogonalization matrix for the instantaneous basis function vector
$\boldsymbol{\Sigma}_0$	Singular values of $\boldsymbol{\Psi}_0$
$sinr_{\mathrm{FD}}$	SINR of an IBFD communications system in general
$sinr_{\mathrm{DC}}$	SINR after digital cancellation
$sinr_i^{\mathrm{d,FD}}$	SINR of the ith DL signal (full-duplex)
$sinr_{\mathrm{B}}^{\mathrm{d,FD}}$	SINR of the backhaul data streams received by the AN (full-duplex)
$sinr_j^{\mathrm{u,FD}}$	SINR of the jth UL signal (full-duplex)
$sinr_{\mathrm{B}}^{\mathrm{u,FD}}$	SINR of the backhaul data streams received by the BN (full-duplex)
$sinr_i^{\mathrm{d,HD}}$	SINR of the ith DL signal (half-duplex)

$sinr_{\mathrm{B}}^{\mathrm{d,HD}}$	SINR of the backhaul data streams received by the AN (half-duplex)
$sinr_{j}^{\mathrm{u,HD}}$	SINR of the jth UL signal (half-duplex)
$sinr_{\mathrm{B}}^{\mathrm{u,HD}}$	SINR of the backhaul data streams received by the BN (half-duplex)
$sinr_{i}^{\mathrm{d,RL}}$	SINR of the ith DL signal (hybrid relay)
$sinr_{\mathrm{B}}^{\mathrm{d,RL}}$	SINR of the backhaul data streams received by the AN (hybrid relay)
$sinr_{j}^{\mathrm{u,RL}}$	SINR of the jth UL signal (hybrid relay)
$sinr_{\mathrm{B}}^{\mathrm{u,RL}}$	SINR of the backhaul data streams received by the BN (hybrid relay)
$sinr_{\mathrm{HD}}$	SINR of a HD communications system in general
$sqnr$	SQNR of the ADC in linear power units
$s(t)$	Received signal of interest
t	Time variable
T_{s}	Sampling period
U	Number of UL UEs
$\mathbf{U}_0$	Left singular vectors of $\mathbf{\Psi}_0$
$\mathbf{U}_{\boldsymbol{\psi}}$	Matrix containing the eigenvectors of the covariance matrix $\mathbf{R}_{\boldsymbol{\psi}}$
$\mathbf{V}_0$	Right singular vectors of $\mathbf{\Psi}_0$
V_{ADC}	Peak-to-peak voltage range of the ADC
W	Operating bandwidth
$w_{i,\mathrm{RFC}}$	Coefficient of the ith RF canceller tap
$x(n)$	The original transmit signal in a SISO transceiver
$x_j(n)$	The original jth transmit signal
$x_{\mathrm{PA}}^{\mathrm{IMD}}(t)$	Nonlinear distortion produced by the TX PA
$x_{\mathrm{PA}}^{\mathrm{in}}$	Input signal of the TX PA
$x_{\mathrm{IQ}}^{\mathrm{TX}}(t)$	Signal at the TX I/Q mixer output
$x_{\mathrm{PA}}(t)$	Output signal of the transmitter in a SISO transceiver
$x_{j,\mathrm{PA}}(t)$	Output signal of the jth transmitter
$y(t)$	Total received signal in a SISO transceiver
$y_{\mathrm{ADC}}(n)$	Signal after the ADC in a SISO transceiver
$y_{i,\mathrm{ADC}}(n)$	Signal after the ADC in the ith receiver
$\mathbf{y}_{i,\mathrm{ADC}}$	Signal vector after the ADC in the ith receiver
$y_{\mathrm{DC}}(n)$	Signal after the digital canceller in a SISO transceiver
$y_{i,\mathrm{DC}}(n)$	Signal after the digital canceller in the ith receiver
$y_{\mathrm{LNA}}^{\mathrm{IMD}}(t)$	Nonlinear distortion produced by the RX LNA
$y_{\mathrm{MX}}^{\mathrm{IMD}}(t)$	Nonlinear distortion produced by the RX I/Q mixer
$y_{\mathrm{VGA}}^{\mathrm{IMD}}(t)$	Nonlinear distortion produced by the RX VGA
$y_{\mathrm{IQ}}^{\mathrm{RX}}(t)$	Signal at the RX I/Q mixer output
$y_{\mathrm{LNA}}(n)$	Signal after the RX LNA
$y_{\mathrm{RFC}}(t)$	Signal after RF cancellation in a SISO transceiver

$y_{i,\mathrm{RFC}}(t)$	Signal after RF cancellation in the ith receiver
$z_{\mathrm{q}}(n)$	Quantization noise produced during the analog-to-digital conversion
$z(t)$	Total noise signal
$z_{\mathrm{RFC}}(t)$	Noise produced by the RF canceller
$z_{\mathrm{rx}}(t)$	Additional noise produced in the receiver
$z_{\mathrm{th}}(t)$	Thermal noise signal
$z_{\mathrm{tx}}(t)$	Additional noise produced in the transmitter
α_{AN}	Total amount of SI suppression in the AN.
α_{BN}	Total amount of SI suppression in the BN.
β_{3dB}	3-dB bandwidth of the phase noise process.
γ_{d}	Auxiliary variable for DL-related quantities
γ_{u}	Auxiliary variable for UL-related quantities
$\gamma_{\mathrm{r}}^{\mathrm{B}}$	Auxiliary variable for quantities related to backhauling DL data
$\gamma_{\mathrm{t}}^{\mathrm{B}}$	Auxiliary variable for quantities related to backhauling UL data
γ^{B}	Auxiliary variable combining $\gamma_{\mathrm{t}}^{\mathrm{B}}$ and $\gamma_{\mathrm{r}}^{\mathrm{B}}$ with α_{AN} and α_{BN}
η	Duplexing parameter (proportion of time spent in the DL timeslot)
θ^{FD}	Auxiliary variable determining the feasibility of the full-duplex scheme
$\mathbf{\Lambda_\Psi}$	Diagonal matrix containing the eigenvalues of $\mathbf{R}_{\boldsymbol{\psi}}$
$\mathbf{M}$	Diagonal matrix containing the step sizes of all the basis functions
$\boldsymbol{\xi}$	Exp. vector for the nonlin. signal model with crosstalk and I/Q imb.
ρ_{d}	Per-UE DL data rate requirement
ρ_{u}	Per-UE UL data rate requirement
σ_{n}^2	Receiver noise floor
τ_{PN}	Delay between the TX and RX I/Q mixers
τ_i	Delay of the ith RF canceller tap
$\phi(t)$	Phase noise of the shared TX/RX LO
$\phi_{\mathrm{eff}}(t)$	The effective phase noise of the direct SI term with a shared LO
$\phi_{\mathrm{rx}}(t)$	Phase noise of the RX LO
$\phi_{\mathrm{tx}}(t)$	Phase noise of the TX LO
$\psi_p(\cdot)$	Static pth-order nonlinear basis function
$\psi_{\boldsymbol{\xi}^k}(\cdot)$	Static basis function corresponding to exponent vector $\boldsymbol{\xi}^k$
$\boldsymbol{\psi}_{\mathrm{L}}(n)$	Instantaneous basis function vector (linear model)
$\boldsymbol{\psi}_{\mathrm{WL}}(n)$	Instantaneous basis function vector (widely linear model)
$\boldsymbol{\psi}_{\mathrm{NL}}(n)$	Instantaneous basis function vector (nonlinear model)
$\boldsymbol{\psi}_{\mathrm{NLC}}(n)$	Inst. basis funct. vector (nonlin. model with crosstalk and I/Q imb.)
$\boldsymbol{\psi}(n)$	Generic instantaneous basis function vector
$\widetilde{\boldsymbol{\psi}}(n)$	Orthogonalized instantaneous basis function vector

$\mathbf{\Psi}(n)$	Complete basis function vector
$\widetilde{\mathbf{\Psi}}(n)$	Complete orthogonalized basis function vector
$\breve{\mathbf{\Psi}}(n)$	Reduced basis function vector
$\mathbf{\Psi}$	Convolution data matrix
$\mathbf{\Psi}_0$	Convolution data matrix weighted with the initial channel estimate

CHAPTER 1

Introduction

1.1 Background and Motivation

Considering the immense growth in the amount of wirelessly transferred data over the past decades, there is little doubt that wireless communications will continue to be one of the cornerstones of the developed world also in the future. More and more information is transferred over-the-air every day, calling for increased efficiency from the wireless communications systems in every aspect. However, as there is a physical limit on how much information can be transferred with radio waves, it is imminent that at some point in the future the ever-increasing data transfer requirements can no more be fulfilled with the existing technologies. In other words, there will be a time when all the usable spectrum has been utilized as efficiently as physically possible, and only a truly radical innovation can increase the wireless data transfer capability any further. Already now, the currently implemented systems are incapable of obtaining the extremely high data rate targets specified for the future fifth generation (5G) networks, mainly due to the congestion of the radio spectrum [67, 98, 177]. Thus, new and innovative technologies must be invented for reaching the goals of 5G and thereby ensuring that the wireless data transfer capability keeps on increasing and fulfilling the demands of the modern society [91, 183, 204, 237].

To this end, several techniques and solutions have been proposed by the research community, such as increasing the amount of antennas to the order of tens or even hundreds [31, 138, 161, 203], or simply communicating on a higher center frequency where bandwidth is abundant for the time being [23, 25, 77, 202, 264]. These methods will indeed provide higher data rates and improved spectral efficiency, but there is an even more fundamental aspect for truly pushing the boundaries of physical-layer spectral efficiency that is yet to be utilized: wireless inband full-duplex (IBFD) communications [35, 44, 63, 204, 266]. What IBFD refers to is simultaneously transmitting and receiving radio signals on the *same center frequency within the same device*. Considering the fact that practically all the current systems operate in a half-duplex (HD) manner, dividing transmission and reception within the device either in time with time-division duplex (TDD) or in frequency with frequency-division duplex (FDD), IBFD-capable radios can

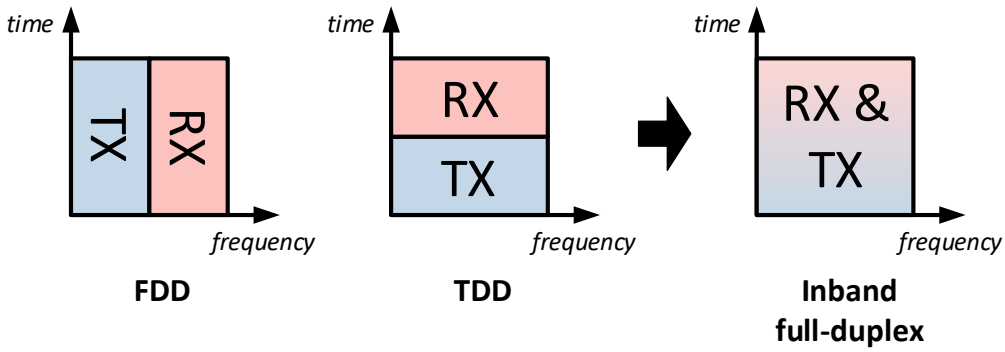


Figure 1.1: A simplified illustration regarding the usage of the spectral and temporal resources in TDD, FDD, and IBFD systems

as much as double the spectral efficiency. The reason for this is simply that neither the temporal nor spectral resources need to be shared between transmission and reception, meaning that the whole available time-frequency resource can be used for both. As a result, the effective resources are doubled, as illustrated also in Fig. 1.1.

The increase in the spectral efficiency can also be demonstrated using the Shannon-Hartley theorem, which specifies the relationship between the capacity, bandwidth, and signal-to-interference-plus-noise ratio (SINR) of a given communications system, assuming Gaussian-distributed noise-plus-interference signal [215]. Namely, the two-way capacities of HD and IBFD communications links can be expressed as follows:

$$C_{\mathrm{FD}} = 2\,W \log_2\left(1 + \mathit{sinr}_{\mathrm{FD}}\right),$$
$$C_{\mathrm{HD}} = \quad W \log_2\left(1 + \mathit{sinr}_{\mathrm{HD}}\right),$$

where W is the overall available bandwidth, $\mathit{sinr}_{\mathrm{FD}}$ is the SINR of the IBFD communications system, and $\mathit{sinr}_{\mathrm{HD}}$ is correspondingly the SINR of the HD system. Assuming that $\mathit{sinr}_{\mathrm{FD}} \approx \mathit{sinr}_{\mathrm{HD}}$, the capacity of the IBFD system is indeed two times higher than that of the HD system, even though the same amount of bandwidth is used. In fact, even if $\mathit{sinr}_{\mathrm{FD}} < \mathit{sinr}_{\mathrm{HD}}$, the IBFD communications system is still likely to outperform the HD system as the SINR affects the capacity inside the logarithm, while the factor two is outside the logarithm. Considering the extreme scarcity of bandwidth, especially in the ultra high frequency (UHF) band, providing such an increase in the spectral efficiency is highly valuable.

However, the inherent challenge of wireless IBFD communications is the problem of self-interference (SI). Namely, any device transmitting and receiving simultaneously on the same frequency band will produce extremely powerful interference to its own receiver (RX) chain, commonly referred to as SI. Moreover, unlike in FDD systems where transmission and reception occur on different frequency bands with a wide separation, in IBFD radios the own transmission cannot obviously be filtered out with a duplexer. If not properly managed, this SI greatly reduces the SINR of the received signal in an IBFD transceiver when compared to a HD transceiver, i.e., resulting in $\mathit{sinr}_{\mathrm{FD}} \ll \mathit{sinr}_{\mathrm{HD}}$

when referring to the above discussion. Consequently, SI can significantly reduce or even completely abolish the obtainable capacity gain of IBFD. Therefore, the central research challenge for IBFD systems is to develop methods and techniques for canceling the SI by some means. Moreover, the accuracy of the SI cancellation solutions must be extremely high since the signal emitted from the own transmitter (TX) chain can easily be over 100 dB stronger than the desired signal of interest [**P1**], [204]. Hence, in such a case, the power of the SI must be attenuated to a *tenth of a billionth* of its original power to even achieve a low SINR of 0 dB. Keeping in mind that the SI fully overlaps with the signal of interest in the frequency domain and therefore no filters can be used, this is a truly daunting challenge.

In principle, however, the SI can be cancelled rather easily: since the transceiver obviously knows its own transmit signal, it can simply subtract it from the received signal. Assuming that the possible channel effects up to the point of subtraction are known, the SI could in fact be perfectly cancelled with this simple principle. What makes SI cancellation challenging in reality is obtaining sufficiently accurate knowledge about the overall coupling channel, i.e., knowing exactly how the SI signal is distorted while propagating from the transmitter to the receiver. In particular, while the effects of the wireless coupling channel between the transmitter and the receiver can be compensated for in a relatively straightforward matter, in many cases the SI signal is distorted also by the TX and RX circuitry. Such distortion, resulting from various analog impairments within the transceiver, cannot usually be captured by the same models that apply to wireless propagation, thereby making accurate SI cancellation rather cumbersome.

What is more, even a relatively insignificant level of unknown distortion in the SI waveform can be a very powerful source of interference for a weak received signal of interest. For instance, the strictest considered error vector magnitude (EVM) for a long term evolution (LTE) base station (BS) is specified as 3.5% [68], which basically means that the total distortion component within the transmitted signal must be 29 dB weaker than the transmit signal itself. This stems from the fact that it is not economically sensible to manufacture devices that outperform the requirements of the specifications by a large margin. In the context of IBFD transceivers, this means that the SI can be cancelled only by 29 dB without any knowledge regarding the dominant distortion components in the transmitted signal. In this example, assuming a typical BS transmit power of 30 dBm and 30 dB of physical isolation between the transmitter and the receiver, the residual SI power would be roughly -30 dBm, rendering any practical receiver incapable of proper operation.

Hence, the first central research challenge for making wireless IBFD communications a reality is the *modeling* and *canceling* of the SI, both of which require extremely high accuracy. In the simplest scenario, the SI model consists only of the physical propagation channel between the transmitter and the receiver, meaning that the observed SI is assumed to be a linear combination of differently delayed *multipath components* of the known transmit signal. However, as discussed above, in such a case the EVM of the transmitter typically limits the amount of obtainable SI cancellation, and consequently the power of the residual SI is too high with any reasonable transmit power. Therefore, facilitating IBFD communications in a practical transceiver requires more comprehensive models for the SI signal that incorporate also the analog impairments that in practice set the EVM. This will facilitate the cancellation of SI beyond the limit established

by the EVM, potentially down to the level of the RX noise floor if all the significant impairments are taken into account when modeling and canceling the SI.

In this thesis, the emphasis is on the SI cancellation performed in the *digital domain*, that is, after the analog-to-digital converter (ADC) of the receiver chain. As will be discussed in more detail in Chapter 2, the purpose of digital SI cancellation is to fully suppress the residual SI that still remains after the different analog or radio frequency (RF) cancellation schemes. The benefit of digital domain cancellation is the increased flexibility in terms of modeling and parameter estimation, which facilitates the use of advanced SI signal models. This means that the significant analog impairments can be explicitly included in the modeling within the digital canceller, and consequently they do not pose a limit for the cancellation performance, as discussed above. Considering that in many cases the EVM of the transmitter is dominated by the nonlinear behavior of its power amplifier (PA) **[P1, 18]**, incorporating a model of this nonlinearity source into the digital canceller can provide a significant improvement in the digital cancellation performance **[P1, P4, P5]**.

In fact, combining such nonlinear digital cancellation with different analog/RF cancellation techniques has been shown to suppress the SI nearly perfectly **[P3–P5]**, [28, 50], meaning that IBFD communications is indeed possible also in reality. Namely, the works in **[P3, P4]**, [28, 50] cancel the SI to the level of the receiver noise floor with two active cancellation stages, one in the RF domain and one in the digital domain, while the work in **[P5]** achieves this by utilizing only digital domain cancellation.[1]

The second central research challenge related to IBFD communications, motivated by the promising performance of these prototypes, is the system-level analysis of networks consisting of IBFD-capable nodes. This type of research is needed in order to determine how to best take advantage of the IBFD capability of the transceivers on a broader scale. Namely, it is not a trivial matter to design a network that actually obtains the promised throughput gains of the IBFD technology. For instance, the network can consist of only IBFD capable devices [40, 252, 253, 271], or it could have an IBFD BS serving legacy HD mobile users **[P7, 134]**, [46, 60, 73]. The former has the benefit of greater flexibility since all devices are capable of engaging in bidirectional IBFD data transfer, while the latter can utilize cheaper hardware in the mobile terminals. The research into the optimal IBFD network architecture is still largely ongoing, and hence there is no answer available regarding which of these, if either, is the best solution. In this thesis, the system-level analysis mainly concentrates on the architecture where only the BS, or access node (AN), is IBFD capable, while the findings regarding SI cancellation obviously make no assumption regarding the network where the devices might be utilized.

In general, implementing commercially feasible IBFD radio transceivers can be considered an important part in utilizing the available spectrum as efficiently as possible. The stringent requirements on SI cancellation obviously require some additional hardware and computational resources within such an IBFD transceiver, but taking into account that radio spectrum is a limited natural resource, this is most likely an acceptable price for the increased spectral efficiency. For this reason, scientific research into the aforementioned aspects of IBFD communications systems is highly meaningful as it paves the way for the potential commercial implementations of the future.

[1] These prototypes are discussed in more detail in Chapter 5.

1.2 Objectives and Scope of the Thesis

The basic research problem or question behind the work presented in this thesis is whether wireless IBFD communications is possible in practice or not. Namely, as already discussed above, operating in IBFD mode is trivial for an ideal communications system but many challenges arise when considering the practical limitations of modern radio transceivers. Hence, the objective of this thesis is to investigate the possibility of IBFD communications when taking also all the practical aspects into account. What is more, in addition to merely studying whether IBFD communications is possible or not, another important objective of this thesis is to also demonstrate true IBFD operation by developing advanced digital-domain techniques for SI cancellation. These proposed solutions are then evaluated with real-life measurements involving an actual IBFD prototype. The obtained results demonstrate that, by incorporating some of the digital SI cancellation algorithms presented in this thesis, IBFD communications is possible with real radio transceivers.

In addition to the practical aspects behind IBFD transceivers and SI cancellation, another research question investigated in this thesis is how to best utilize IBFD communications on a *system level*, taking into consideration the various sources of interference incurred by the IBFD operation. In particular, this thesis analyzes the general feasibility of a network where an IBFD-capable AN is serving HD mobile users while also maintaining a wireless backhaul link. Since all of the transmissions are performed on the same frequency band and at the same time, careful power allocation is needed to obtain the required user data rates while also minimizing the power of the interference signals.

Together, these objectives and research topics provide various findings regarding the feasibility and system-level applications of the IBFD technology. In other words, this thesis demonstrates that it is indeed possible to cancel the SI such that IBFD operation is possible while also proving the benefits of utilizing IBFD transceivers in a network.

1.3 Thesis Contributions and Structure

The main contributions of this thesis are the following.

- The most significant analog impairments in terms of the digital domain modeling of the SI waveform are determined by using realistic models for the different transceiver components [**P1–P3, 18, 82, 121, 123, 132, 133, 233**]. Especially, it is shown that the nonlinear distortion of the TX PA and the in-phase/quadrature (I/Q) imbalance are the prevalent impairments under practical circumstances.
- Three different signal models for the residual SI in the digital domain are developed, taking into account either some or all of the significant analog imperfections [**P2–P6, 4, 17, 18, 120–126, 133**]. In addition, also a linear signal model, corresponding to a scenario with ideal transceiver circuitry, is presented for reference. For generality, all of the signal models are derived for a multiple-input and multiple-output (MIMO) IBFD transceiver. Altogether, to the best of the author's knowledge, this thesis presents some of the most comprehensive models for the SI signal in the literature.

- A complete digital SI cancellation solution, utilizing the above signal models, is proposed. It incorporates one of two alternative parameter estimation algorithms for learning the coefficients of the used signal model, both of them also presented in this thesis [**P2–P6, 4, 17, 18, 120–128, 133**]. One of the parameter estimation schemes is based on block-wise processing (least squares), while the other one is an adaptive estimation algorithm (least mean squares).
- A complexity reduction scheme for the signal models, which can be used to reduce the number of parameters, is derived [**P6**]. It is based on principal component analysis (PCA), and it is shown to significantly reduce the number of parameters required for accurately regenerating the SI signal.
- Comprehensive simulation and measurement results evaluating the performance of the proposed digital cancellation solutions are presented [**P1–P6, 4, 17, 18, 82, 120–126, 133, 238**]. The measurements incorporate also different analog SI suppression techniques. These results show that the developed digital SI cancellers are capable of suppressing the SI to the level of the receiver noise floor, thereby facilitating true IBFD operation. In fact, in some of the measurements, more than 100 dB of SI cancellation is obtained which is one of the highest amounts of cancellation reported in the existing literature.
- The optimal transmit power allocation is derived for a network where an IBFD-capable AN is using the same time-frequency resource for both serving HD mobile users and backhauling itself wirelessly [**P7, 129–131, 134**]. The optimal transmit powers are given in closed form, alongside with the condition for the feasibility of any given Quality of Service (QoS) requirements. The proposed IBFD AN is also compared to two alternative solutions where either some or all of the transmissions are multiplexed in time domain. In general, the findings indicate that utilizing the IBFD capability of the AN is greatly beneficial for the considered system.

Full account of the contributions is given in [**P1–P7, 4, 17, 18, 82, 99, 120–134, 160, 193, 194, 233, 238**], while only the essential information and results are collected in this thesis summary. Moreover, to ensure better consistency, the notation used in the thesis summary is slightly different than that used in the publications.

The thesis is organized as follows. Chapter 2 gives an overview regarding the history of IBFD communications while also discussing the recent research on a general level. It also describes the basic operating principles of IBFD transceivers, alongside with providing the baseband-equivalent models of the essential analog impairments. Then, Chapter 3 derives the power levels of the different distortion components present within an IBFD transceiver, which can be used as the starting point for developing the signal models for digital cancellation. After this, Chapter 4 presents four different signal models that can be used for digital modeling and cancellation of the residual SI under analog impairments. Also two alternative parameter estimation schemes are proposed, alongside with a complexity reduction scheme. These signal models are then evaluated in Chapter 5, where the digital cancellation performance is assessed both with simulations and actual RF measurements. Following this, Chapter 6 analyzes a potential network-level application for an IBFD-capable device where the same time-frequency resource is used for uplink (UL), downlink (DL), as well as for wireless backhauling. Finally, the conclusions and potential future research directions are presented in Chapter 7.

1.4 Author's Contributions to the Publications

Originally, this research area was proposed by Prof. Mikko Valkama, and he has also contributed to all of the publications by sharing his thoughts on the contents and different problems while also providing extensive feedback for the manuscript drafts.

In [**P1, P2, P4, P6, P7**], the author of this thesis (later: the Author) did all the mathematical derivations and performed the simulations and/or the digital cancellation of the measured data. In [**P1**], Prof. Markku Renfors provided some guidance in the very early stages of the research, while the other authors helped to polish the contents of the manuscript. In [**P2**], D.Sc. Lauri Anttila shared his ideas especially regarding the estimation bias caused by the PA-induced nonlinear distortion. Moreover, the measurements reported in [**P4**] were for the most part carried out by research assistants under the supervision of the Author, while the digital cancellation and the post-processing were done by the Author himself. In [**P6**], the initial idea was developed together with D.Sc. Lauri Anttila, who also first derived a simplified version of the signal model in [**17**]. Furthermore, the initial ideas behind the analysis presented in [**P7, 134**] were developed together with Prof. Ashutosh Sabharwal and D.Sc. Taneli Riihonen, although all the research work was done by the Author.

Moreover, the work in [**P3, P5**] was done in close collaboration with project partners from Aalto University, who also provided the high-isolation relay antenna used in both of these publications. Especially, in [**P3**], the Author did the analysis of the analog imperfections, alongside with digitally processing the measurement results. In [**P5**], the Author performed the digital cancellation and post-processing of the measured data, while the results regarding the antenna isolation were provided by Mikko Heino.

In addition to the work published in journals [**P1**–**P7**], the Author has also published and co-authored numerous conference articles regarding this topic [**4, 17, 18, 99, 120**–**133, 160, 194, 238**], alongside with participating in writing the journal articles in [**193, 233**] and the book chapter in [**82**]. In addition, one journal article is yet to be officially published [**134**]. Although these publications are not attached into the thesis, they are still part of the Author's overall contribution and have been used in compiling this summary, as referenced throughout the text.

In particular, the paper in [**132**] is essentially a subset of the results presented in [**P1**], with the Author having done the derivations and the writing. Moreover, the work in [**133**] is a joint article with collaboration partners from Aalto University, for which the Author derived the proposed digital cancellation procedure that incorporates also the RX-induced nonlinearities into the modeling. In [**18**], a nonlinear digital cancellation algorithm is derived, with the Author generating the results for the system calculations example, alongside with the simulation results. The work in [**121**], which presents four different signal models to be used in the digital canceler, was conducted solely by the Author, with the co-authors providing feedback regarding the editorial aspects of the paper. The same also applies to [**122**] and [**123**]; in the former, the impact of the received signal of interest on the SI channel estimation is analyzed, while the latter investigates a reference receiver–based digital canceller. The work in [**17**] presents a joint signal model for both the PA-induced nonlinearities and the I/Q imbalance, with D.Sc. Lauri Anttila having done the derivations and the Author having generated the simulation results.

Furthermore, the works in [**127, 128**] analyze the optimal way of estimating the SI channel in terms of maximizing the overall rate region, with [**128**] considering a bidirectional data link and [**127**] investigating a relaying scenario. The initial idea for these publications was developed together with D.Sc. Taneli Riihonen, while the Author derived the results, performed the simulations, and took the main responsibility in writing the papers. In [**99**], a two-tap RF canceller is presented, and the Author mainly contributed to the editorial aspects of this paper. Then, in [**126**] the digital canceller derived in [**17**] is evaluated with RF measurements, with the Author performing the digital cancellation and also writing the paper. The work in [**125**], on the other hand, presents the overall cancellation performance for the same single-antenna IBFD prototype discussed also in [**P4**], although with further technical details. There, the digital cancellation of the measurement data was performed by the Author, alongside with the actual writing of the manuscript.

In [**124**], a nonlinear digital canceller is evaluated for asymmetric TX and RX frequencies such that the SI is cancelled over a wider bandwidth than actually occupied by the transmit signal. Also this paper was completely composed and written by the Author. The work in [**238**] provides a detailed study of a three-tap RF canceller with a digital control algorithm, the Author mainly providing supervision with regard to the writing of the paper. An electrical balance duplexer–based IBFD transceiver architecture is then analyzed and evaluated in [**160**], where the Author helped in developing the control algorithm and editing the paper. Moreover, in [**120**], a real-time implementation of the nonlinear digital canceller reported in [**125**] is presented and evaluated. The Author's responsibility here was to write the paper and provide guidance on the implementation and measuring of the real-time digital canceller. Another implementation of the nonlinear digital canceller is reported in [**4**], where the performance of different multi-core processing platforms is analyzed in the context of digital SI cancellation. For this work, the Author provided help in implementing the algorithm while also proofreading and editing the paper.

The work in [**129**] derives the optimal transmit power allocation for a network with a self-backhauling IBFD AN under the objective of maximizing the sum-rate. Here, the Author derived the results and wrote the paper, while the co-authors provided guidance. Furthermore, in [**134**], the sum-rate of the same system is numerically optimized, albeit using a somewhat more comprehensive model. The Author was responsible also for composing this paper. A similar system is considered also in [**131**], although there the objective is to minimize the transmit powers under given minimum data rate requirements. Again, the Author derived the results and wrote the paper under the guidance of the co-authors. Moreover, the fundamental feasibility of this type of a network with a self-backhauling AN is analyzed in [**130**], albeit for a simplified scenario in comparison to [**P7**]. Also the results for this paper were derived by the Author, alongside with the actual writing of the work.

In addition, the Author participated also in the writing of the journal article in [**233**] by providing feedback on different technical aspects during the writing process. Furthermore, the book chapter in [**82**] contains also some contributions from the Author, most prominently the discussion regarding the analog impairments, the different digital cancellation solutions, and the obtained results. Recently, the Author has also applied some of the contributions of this thesis to military communications [**193, 194**].

1.5 Nomenclature

In this thesis, all the signals are assumed to be zero-mean and complex-valued. In other words, baseband-equivalent modeling is utilized throughout the thesis, with the signals being of the form $x(t) = x_{\mathrm{I}}(t) + \jmath x_{\mathrm{Q}}(t)$, where the continuous-time variable is denoted by t, $x_{\mathrm{I}}(t)$ is the in-phase component, $x_{\mathrm{Q}}(t)$ is the quadrature-phase component, and $\jmath$ is the imaginary unit with the property $\jmath^2 = -1$. The corresponding RF signal is then simply $x_{\mathrm{RF}}(t) = \mathrm{Re}\{x(t)e^{\jmath\omega_c t}\}$, where $\mathrm{Re}\{\cdot\}$ gives the real part of the signal, e is Euler's number, and ω_c is the angular carrier frequency. For discrete-time signals, the sample index is denoted by n, and the corresponding digital baseband signal is consequently $x(n) = x_{\mathrm{I}}(n) + \jmath x_{\mathrm{Q}}(n)$. The absolute value of both real and complex numbers is denoted by $|\cdot|$, while the complex conjugate is written as $(\cdot)^*$. The Fourier transform is denoted by $\mathcal{F}\{\cdot\}$, with f being used as the frequency variable. Furthermore, the convolution is denoted by $\star$, and it is defined for continuous-time signals as follows [104, p. 443]:

$$x(t) \star y(t) = \int_{-\infty}^{\infty} x(\tau)y(t-\tau)\,\mathrm{d}\tau.$$

Correspondingly, the convolution of two discrete-time signals is defined as [104, p. 525]

$$x(n) \star y(n) = \sum_{m=-\infty}^{\infty} x(m)y(n-m).$$

Considering then the used vector/matrix notations, the vectors are denoted by bold and lower case variables (e.g., $\mathbf{x}$), while the matrices are correspondingly bold and upper case variables (e.g., $\mathbf{X}$). Furthermore, the transpose and the Hermitian transpose are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively, while the element on the ith row and jth column of a matrix $\mathbf{X}$ is expressed as $\{\mathbf{X}\}_{ij}$. Similarly, the ith element of a vector $\mathbf{x}$ is written as $\{\mathbf{x}\}_i$, while its L^1-norm is defined as $\|\mathbf{x}\|_1 = \sum_i |\{\mathbf{x}\}_i|$. The Hadamard product between two matrices of same dimensions is denoted by $\circ$. In addition, the statistical expectation is denoted by $\mathbb{E}[\cdot]$ for both scalars and vectors/matrices.

CHAPTER 2

Inband Full-Duplex: Basic Principles and Essential System Models

This chapter provides essential background information regarding IBFD communications, including the historical developments and research activities, the basic concept in some detail, and the most typical system models for IBFD devices. Also the baseband-equivalent models for the significant analog imperfections are provided. These are then used later as a basis for analyzing the analog impairments, as well as for developing the digital cancellation algorithms.

2.1 History and Early Developments

Simultaneous transmission and reception (STAR) on the same frequency band within a single device is in fact a rather old technique as it has been applied in continuous wave (CW) radars ever since the 1940s [157, 204, 227, 228]. There, in order to detect the targets, the radar must be capable of receiving the echoes from its own transmission while also continuously transmitting itself. What makes this problematic is the fact that such a CW radar mainly receives its own transmitted signal, which is orders of magnitudes stronger than the echoes from the possible targets since the transmit signal is coupling directly from the device itself. This is somewhat equivalent to the problem of SI in the context of generic IBFD transceivers. In the early CW radars, the necessary SI attenuation was achieved by means of antenna isolation if using separate TX and RX antennas [228], or with a circulator in a single-antenna radar [227]. The former is typically referred to as a *bistatic* radar, while the latter is called a *monostatic* system. Perhaps one of the most extreme examples of a bistatic radar is the experiment performed on the Apollo 14, 15, and 16 moon flights, where the command and service module (CSM) transmitted a signal towards the moon, whose reflection was then received in

the earth and analyzed to gain understanding regarding the properties of the lunar surface [93–95, 221]. In this case, the radar transmitter was hence in the lunar orbit, while the receiver was situated in the earth, corresponding to nearly 400 000 km of physical separation between the two. In the more earthly applications, however, both the monostatic and bistatic radars suffered from a relatively low level of isolation between the transmitter and the receiver, which meant that the CW radars were initially limited to rather low transmit powers, and consequently also to low distance targets. Hence, to allow for a longer detection range, it has been more common to utilize so-called pulsed radars, which separate the transmission and reception in time [204, 227].

Since these early systems, there have been several advances in the techniques for suppressing the own transmission in CW radars. In particular, various active solutions have been proposed for canceling the transmitter leakage at the receiver input, allowing for an extended detection range. One of the earliest proposals was the so-called feed-through nulling, which is essentially implementing a closed-loop RF canceller where part of the transmit signal is subtracted from the received signal [178]. In order to actually cancel the leakage, the phase and amplitude of the cancellation signal must be adjusted in real time by using a vector modulator (VM) and closed-loop learning, thereby facilitating accurate tracking of the leakage channel. The implementation in [178] is reported to achieve around 60 dB of cancellation. More recently, a similar and a significantly cheaper solution has been presented in [27], which is reported to attenuate the leakage by 33 dB or more. The research into the different leakage suppression methods in the context of radars has been active up until the recent years, and there are various advanced analog- and digital-domain solutions available [33, 112, 113, 149, 150, 186]. Furthermore, simultaneous radar operation and data transfer, facilitated by IBFD communications, has also been recently considered as a means of transmitting feedback information in bistatic radar systems with distant transmitters and receivers [180].

Another historical issue closely related to IBFD communications is the problem of echo on a telephone line. In particular, mainly due to the various junctions with imperfect impedance matching in the telephone network, the speaker's own transmitted voice typically echoed back and was consequently heard by the speaker [30, p. 17]. With sufficiently long delays, this proved to be extremely disturbing, meaning that a solution was urgently needed. Initially, the problem was alleviated by using so-called telephone hybrids when converting the incoming and outgoing signals between 2-wire and 4-wire circuits. With the help of transformers, the hybrid maintained a certain level of isolation between the incoming and outgoing speech signals, thereby weakening the signal echoes [30, p. 71]. However, the drawback of using a hybrid to cancel the echo is that it is based on impedance matching, whose accuracy is usually rather low under practical circumstances. This limited the amount of achievable echo cancellation heavily and called for more advanced solutions. To this end, AT&T Bell Labs developed and patented a so-called echo canceller in the 1960s and 1970s [111]. What makes this especially significant in terms of IBFD transceivers is that the echo canceller very closely resembles some of the more modern analog multi-tap SI cancellers **[238]**, [28, 119], and consequently can be considered as a predecessor to the modern RF cancellation solutions. For the developed telephone echo canceller, a maximum suppression of 30 dB is reported [111]. Similar type of echo cancellers have also been suggested for asymmetric digital subscriber line (ADSL) modems, where they can facilitate IBFD operation [102]. However, even

though such IBFD mode was included in the International Telecommunication Union (ITU) recommendation G.992.1, it was never commercially implemented on a wide scale since the data rate requirements in the ADSL networks are highly asymmetric. This greatly decreases the gain provided by IBFD operation.

More recently, the IBFD technology has been successfully applied to repeaters or relays [135–137, 155, 175, 210, 218, 224, 225, 230, 261]. This is a well-suited application for IBFD transceivers, since it inherently calls for STAR on the same center frequency. What is more, a relay typically receives from a different direction than it transmits in, allowing for more physical isolation between the TX and RX chains [**P3, P5**]. This somewhat alleviates the requirements on the active SI cancellation solutions, resulting in a simpler implementation [**P5**]. Relaying is of course also possible in a HD manner by dividing the transmission and reception in time or in frequency, but then the network must explicitly consider the existence of the relay by some form of scheduling or frequency planning. On the other hand, if assuming a reasonably small processing delay, an IBFD repeater can be entirely transparent from the network's perspective since it does not require additional temporal or spectral resources. This makes it extremely well suitable for, e.g., gap-filling in general [86, 224, 261], or increasing the coverage in digital video broadcasting (DVB) systems [175, 210].

In the earlier proposals of such IBFD relays, the necessary SI isolation was achieved by spatially separating the transmit and receive antennas [14, 224, 225]. The downside of this are of course the relatively large dimensions required for the device, and hence active SI cancellers for relay applications have also been widely studied, including analog [108, 115, 231] as well as digital solutions [80, 87, 90, 158, 198, 200, 208]. Recently, various beamforming, spatial-domain nulling, and power control solutions have also been proposed [36, 48, 49, 83, 86, 106, 139, 153, 192, 196, 197, 211, 214, 239], alongside with several more information theoretic analyses of IBFD relays in general [54, 107, 151, 152, 199, 220, 249, 259]. Furthermore, IBFD relaying has also been analyzed in the context of energy harvesting and wireless power transfer [170, 250, 262, 270].

In conclusion, regardless of the wide body of literature involving different kinds of SI cancellers for all types of IBFD radios, the idea of actual two-way data transfer over the same time-frequency resource received rather little interest up until the recent years. The likely reason for this is that the amount of cancellation needed for any meaningful data rate and/or communications distance in such a scenario is clearly beyond the capabilities of the solutions developed for radars and echo cancellers [204]. Realizing such immense isolation levels for a generic radio device was simply considered too steep a price for a mere twofold improvement in the spectral efficiency during a time when spectrum was abundant. Even so, there are still some patents for somewhat more generic IBFD transceivers dating back to the early 1980s [246, 247], indicating that their potential in data transfer applications has still been understood, although the necessary technology and commercial interest has been lacking. Eventually, however, the heavy congestion of the favorable parts of the radio spectrum made the cost of SI cancellation comparable to the benefits of doubling the spectral efficiency, thereby motivating the research also into IBFD data transfer applications.

2.2 Modern Take on Full-Duplex: Bidirectional Data Transfer over the Same Time-Frequency Resource

Due to the aforementioned aspects, the interest in wireless IBFD communications as a duplexing method in bidirectional data transfer applications has been constantly increasing in the more recent years [62, 188, 204, 266]. Starting from [35], where some of the first experiments regarding bidirectional IBFD communications are reported, the research has been intensive. To first quantify the required amount of SI suppression in these generic IBFD transceivers, let us assume a maximum LTE user equipment (UE) transmit power of 23 dBm, and a sensitivity requirement of −90 dBm [69]. With these example system parameters, the SI should be cancelled in total by 113 dB to attenuate it to the level of the receiver noise floor, where it would still decrease the sensitivity by 3 dB. Although there is usually some physical isolation between the TX and RX chains, several active cancellation stages are clearly needed as it is extremely challenging to achieve such suppression levels by passive isolation mechanisms alone.

Starting from the first IBFD prototype implementations, the prevalent solution has been to utilize two active SI cancellation stages to suppress the SI sufficiently: cancellation in the analog/RF domain and then subsequent cancellation in the digital domain **[P3–P5]**, [3, 19, 28, 29, 44, 59, 61–63, 103, 265]. The analog/RF cancellation is usually performed already before the actual RX chain to limit the total power entering the RX low-noise amplifier (LNA). In principle, the RF canceller subtracts the regenerated SI signal from the overall received signal in the RF domain and thereby ensures that the extremely high SI power does not saturate the LNA or damage any of the delicate components in the RX chain. However, due to the immense SI cancellation requirements, further cancellation is typically required in the digital domain. There, the original baseband transmit data is used to regenerate the residual SI, which is then subtracted from the overall digitized signal to suppress the remaining SI **[P3, P4]**. Complemented with the physical isolation between the transmitter and the receiver, this type of an overall cancellation solution has been shown to be sufficient for suppressing the SI below the receiver noise floor **[P4, P5]**, [28, 29, 50]. It is also possible to omit either the RF or digital canceller but this requires advanced techniques for significantly increasing the amount of passive isolation **[P5]**, [8]. These different cancellation solutions are covered in more detail in Section 2.3 below, while a more detailed review of the different IBFD prototypes and their corresponding SI cancellation capabilities is given in Chapter 5.

Stemming from these successful demonstrator implementations, there is also a wide body of more theoretical and fundamental research into the data transfer–oriented applications of wireless IBFD communications. For instance, the medium access control (MAC)–related aspects of networks consisting of IBFD devices have been widely analyzed, since the IBFD capability obviously affects the whole channel access procedure [12, 43–47, 58, 78, 103, 147, 148, 174, 222, 226, 251, 271]. Hence, new solutions for MAC mechanisms are needed when IBFD transceivers are introduced. As a specific example, utilizing IBFD–capable transceivers in a network based on carrier sense multiple access with collision avoidance (CSMA/CA) naturally removes or alleviates the well-known *hidden node problem* illustrated in Fig. 2.1a, where a collision occurs when two nodes outside each other's transmission range try to transmit data signals to the same receiving node at the same time. In particular, many of the IBFD MAC algorithms solve this

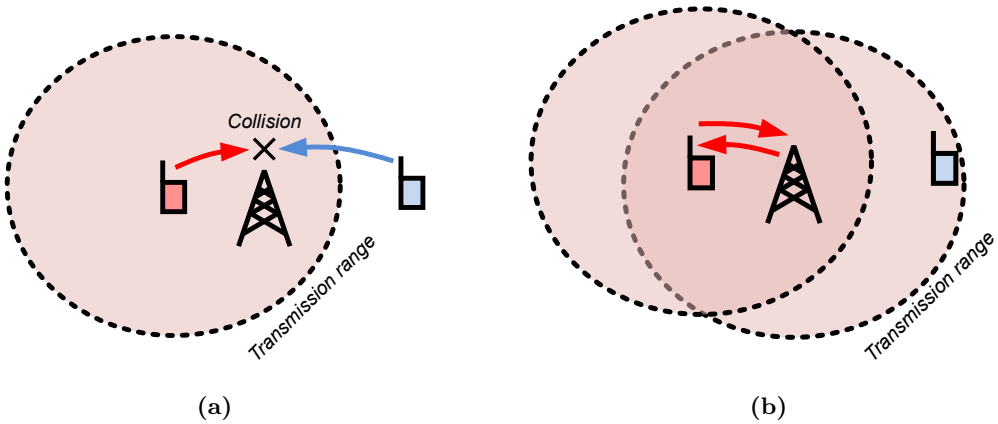


Figure 2.1: (a) The hidden node problem in a CSMA/CA network with HD devices and (b) the same scenario with IBFD-capable devices, where no collision occurs.

issue by forcing also the receiving node to transmit either a data signal or a busy tone, facilitated by its IBFD capability [44, 47, 58, 78, 103, 174, 222, 271]. This type of a scenario is illustrated in Fig. 2.1b, where it can be observed that the channel is now reserved around both parties, ensuring that the collision avoidance mechanisms prevent further transmissions.

Furthermore, in a CSMA/CA-based network, the IBFD or STAR capability can also be used to sense the channel while transmitting, which enables much faster detection of a possible collision [147, 148, 226]. This obviously results in a higher average throughput in the network, illustrating that the IBFD technology can also provide performance improvements in the higher layers, in addition to increasing the physical layer spectral efficiency. It should also be noted that STAR is extremely useful also for cognitive radios, where the secondary users must cease their own transmission upon detecting a primary user initiating a transmission [5, 232, 236]. Being capable of continuously sensing the channel obviously decreases the level of interference that the secondary users cause to the primary users, resulting in more efficient and operational cognitive radio networks.

Moreover, in order to decrease the overall deployment complexity of the IBFD networks, systems where only a subset of the transceivers are capable of IBFD operation have also been widely studied. A popular example of this is the system illustrated in Fig. 2.2 where the BS or the AN is IBFD capable, while the UEs or the clients are legacy HD devices [46, 60, 70, 73, 167, 171, 185, 217, 229, 254, 255, 272]. The benefit of this solution is that the BS can serve the UL and DL UEs simultaneously on the same time-frequency resource, while the potentially costly SI cancellation must be performed only in the BS. However, as also shown in Fig. 2.2, a downside of this type of a network is that the transmissions of the UL UEs will produce interference to the receiving DL UEs, and therefore this so-called inter-user-interference (IUI) must be mitigated by some means. There are already various solutions for addressing the IUI, such as assigning the UL and the DL UEs so that their mutual path losses are maximized [46, 60] or interference alignment [65, 114, 205], indicating that this issue can be alleviated at least to some extent. Altogether, a network where only the BS is IBFD capable is hence an

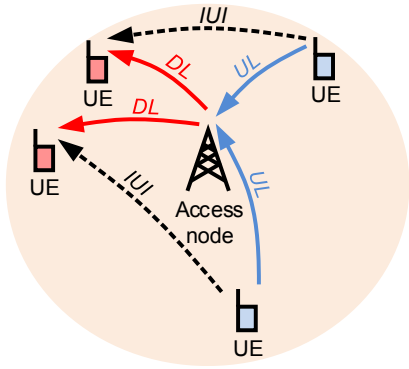


Figure 2.2: An illustration of a network where an IBFD-capable AN is using the same frequency band for simultaneous DL and UL.

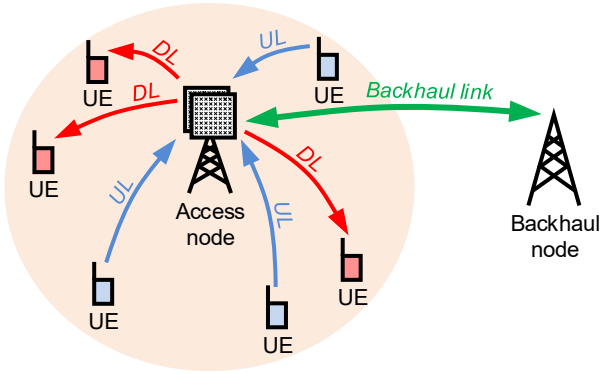


Figure 2.3: An illustration of a network where an IBFD-capable AN is using the same time-frequency resource for DL, UL, and backhauling. Note that this type of a network suffers also from IUI between the DL and UL UEs and the backhaul node, but the different interference links are omitted from this figure to improve its readability.

intriguing commercial application for IBFD transceivers, as it will ensure a lower cost for the UEs while still improving the spectral efficiency via the simultaneous UL and DL transmissions.

A particular example of a network with an IBFD-capable AN is discussed, for instance, in [**P7, 129, 131**], [183, 237] and depicted in Fig. 2.3, where the AN serves the UL and DL UEs simultaneously on the same frequency band while also using the same frequency resources for backhauling the data wirelessly. Such wireless *self-backhauling* is greatly beneficial, for instance, in densely deployed cellular networks, where it would be very expensive to install wired backhaul links for all the individual cells [41, 91, 183, 217, 237, 268]. This type of a situation is a probable scenario in the future 5G networks [67, 98, 110]. As no additional spectral resources are needed for establishing the backhaul connectivity, such a network is spectrally very efficient while requiring no expensive cabling for the backhaul link. However, although omitted from

Fig. 2.3 for a better illustration, the drawback of this type of a scheme are the various interference sources, as the different signals are fully overlapping both in frequency and in time. This calls for careful power allocation in order to minimize the interference powers [**P7**]. Such a network is analyzed in more detail in Chapter 6.

On a more general note, a great deal of attention has also been paid to the obtainable network-level performance gains when utilizing IBFD devices [9, 24, 171, 172, 212, 252, 253, 260, 271]. With a suitable MAC protocol, introducing IBFD transceivers to the network might actually improve the overall throughput by more than 100% [271], although most works report the achievable gains to be less than 100% due to the increased interference levels between the transceivers [212, 252, 253]. Nevertheless, under most circumstances, employing IBFD communications is still shown to provide a higher throughput than the corresponding HD solution [3, 9, 24, 53, 70–72, 171, 172, 212, 260, 271]. Hence, while wireless IBFD communications suffers from the additional cost of SI cancellation, it results in improved spectral efficiency also on a network level. This makes it an intriguing technology for the next generation wireless systems, as long as the cost of the necessary SI cancellers can be kept at a reasonable level.

2.3 Modeling Inband Full-Duplex Transceivers

Let us then consider in more detail the common IBFD transceiver architectures, which facilitate the simultaneous transmission and reception on the same center frequency. In principle, such a transceiver is differentiated from a typical HD transceiver by the SI cancellation stages, having otherwise similar transmitter and receiver structures.

In this thesis, the direct-conversion architecture is considered for both the transmitter and the receiver. This selection is made because direct-conversion transceivers are widely used in the modern wireless systems, thanks to their low power consumption and relative ease of integration on a single chip [2, 169, 190]. As a reference, Fig. 2.4 shows a generic block diagram of such a HD direct-conversion transceiver, illustrating the relevant aspects of the architecture. The defining characteristic of a direct-conversion transceiver is that the signal is downconverted (or upconverted) directly from the RF to the baseband (from the baseband to the RF). This simplifies the transceiver since no intermediate frequencies are used, and some of the amplification and filtering can be done in the baseband [169]. However, as a downside, the direct-conversion architecture suffers from direct current (DC) offset, local oscillator (LO) leakage, and I/Q imbalance, which must be mitigated in the digital domain or by some other means [169, 190]. Especially the I/Q imbalance is a cumbersome issue in the context of IBFD transceivers, as will be discussed in more detail in Section 2.4.1.

To facilitate wireless IBFD operation, Fig. 2.5a illustrates a possible SI cancellation architecture for a direct-conversion transceiver. Firstly, it includes two alternatives for physically isolating the transmitter and the receiver: separate TX and RX antennas, or a shared TX/RX antenna. In the latter, the physical isolation is obtained with a so-called circulator, which is a three-port device where each port is connected directly to the next port while being isolated from the previous port, or vice versa, depending on the direction of rotation (i.e., clockwise or counterclockwise) [184, p. 487]. This means

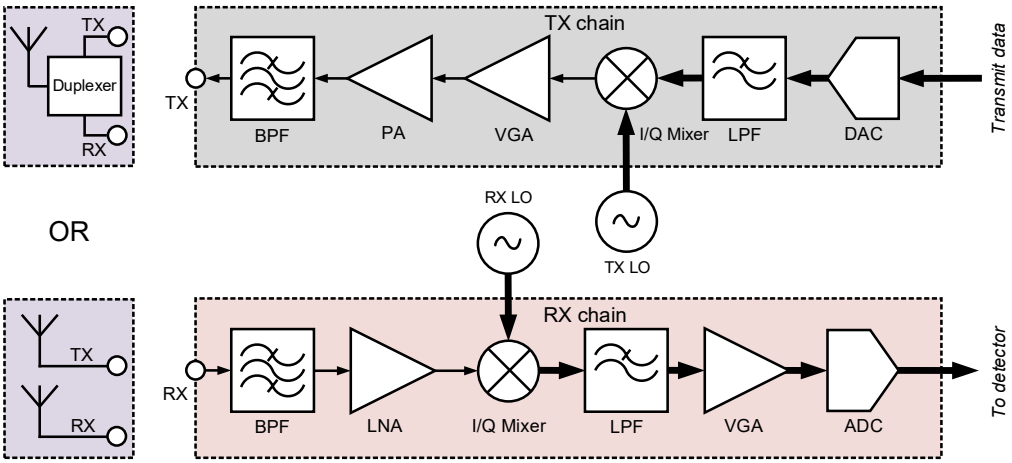


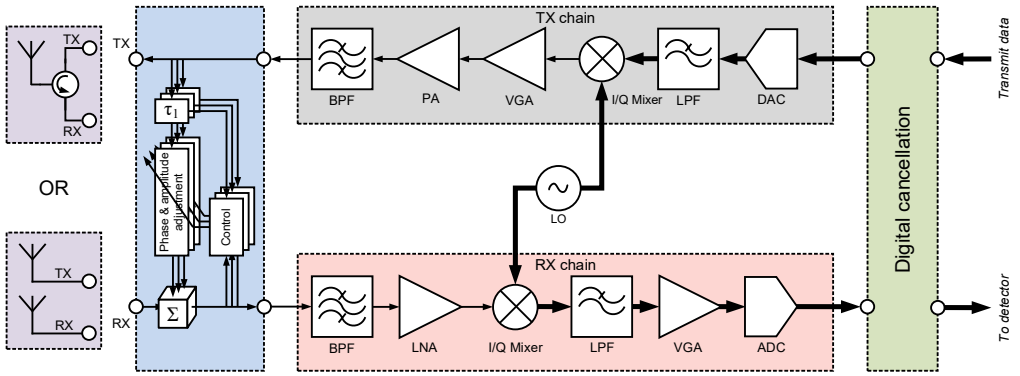
Figure 2.4: An illustration of a HD direct-conversion transceiver, including options for both TDD and FDD operation modes. In the former, either a shared TX/RX antenna or separate TX and RX antennas can be used, while a shared antenna connected to a duplexer is typically used in the latter case. For a full list of the abbreviations, refer to pp. ix–xi.

that the signal propagating in the opposite direction is heavily attenuated, resulting in a certain amount of isolation between the TX and RX chains when using such a circulator in an IBFD transceiver as shown in Fig. 2.5a. The isolation provided by a circulator is typically in the order of 20–40 dB, depending on its size and cost as well as on the used bandwidth, while the attenuation in the desired direction is usually less than half a decibel [**P4**]. It should be noted that the essential modeling of the SI signal is not affected by the adopted antenna architecture, and hence the signal models and algorithms reported in this thesis can be readily applied to both types of systems.

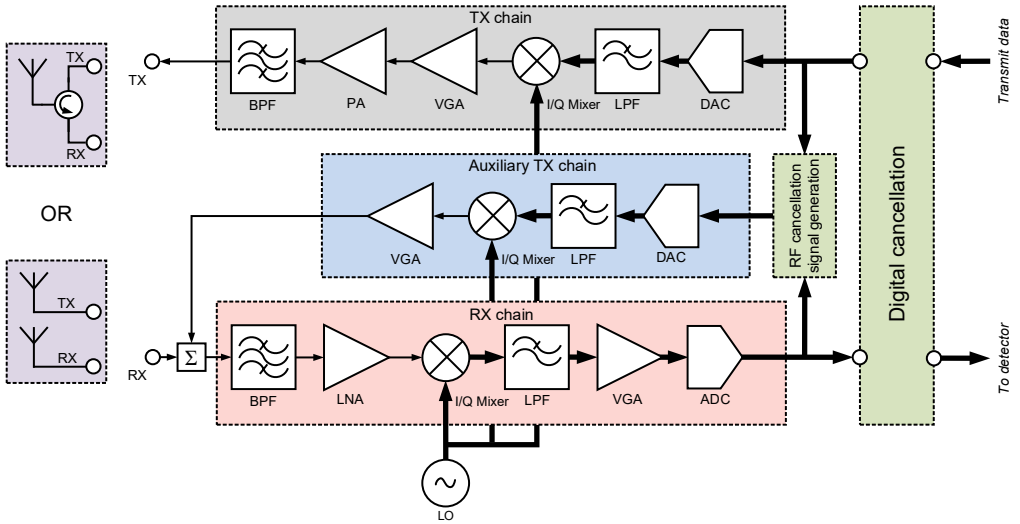
Shared-antenna IBFD operation can also be facilitated by using an electrical balance duplexer (EBD) in place of the circulator. In principle, an EBD utilizes a hybrid transformer to isolate the TX and RX chains connected to the same antenna, the isolation being achieved by mimicking the impedance of the antenna with a tunable balance network [**160**], [56, 140, 141, 263]. Namely, the tunable impedance is used to form a replica of the signal leaking from the TX chain to the RX chain, which is phase shifted by 180° with the hybrid transformer before combining it with the RX signal and consequently suppressing the SI. While an EBD can typically be implemented in a more compact form than a circulator, it requires active tuning to track the antenna impedance, making the fully passive circulator in many respects a more robust option. For these reasons, in this thesis the single-antenna IBFD transceiver is primarily assumed to utilize a circulator, although the essential results in Chapters 3 and 4 can also be applied to an EBD-based architecture.

The overall signal received by an IBFD transceiver, consisting of both the SI and the signal of interest, is then formed as illustrated in Fig. 2.6. With the help of baseband-equivalent modeling, it can be expressed as follows:

$$y(t) = s(t) + h_{\mathrm{SI}}(t) \star x_{\mathrm{PA}}(t) + z(t), \tag{2.1}$$



(a)



(b)

Figure 2.5: (a) The considered direct-conversion IBFD transceiver architecture where RF cancellation is performed using the PA output signal, and (b) the corresponding IBFD transceiver architecture where the RF canceller utilizes an auxiliary transmitter, first proposed in [63]. For a full list of the abbreviations, refer to pp. ix–xi.

where $s(t)$ is the received signal of interest, $h_{\mathrm{SI}}(t)$ is the (wireless) coupling channel between the transmitter and the receiver, $x_{\mathrm{PA}}(t)$ is the transmitter output signal, and $z(t)$ is the overall noise signal.

After this, RF cancellation is performed on the received signal to reduce the SI power entering the actual RX chain. The IBFD transceiver architecture depicted in Fig. 2.5a utilizes a RF cancellation solution where the transmitter PA output signal is used to form the cancellation signal, which is then subtracted from the received signal after proper manipulation. The benefit of this type of a solution is that all the

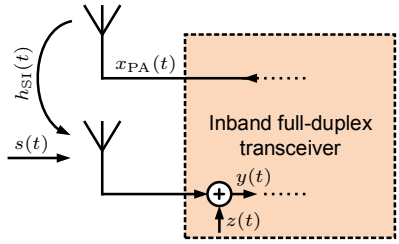


Figure 2.6: The basic model for the total received signal in an IBFD transceiver.

TX-induced impairments are implicitly included in the cancellation signal, and are consequently suppressed by the RF canceller. In principle, this type of an RF canceller aims at modeling and subtracting the strongest SI components, while the multipath reflections are more easily cancelled in the digital domain **[P5]**, [28, 103]. When separate TX and RX antennas are used, there is only one strong SI component present at this stage, i.e., the directly propagating component **[99]**, [59, 119], while the circulator-based architecture results in two powerful SI components: the direct leakage through the circulator and the reflection from the antenna [28, 257]. In the latter case, the power level of the antenna reflection is directly determined by the quality of the antenna matching, and hence this provides a strong motivation to utilize well-matched antennas in single-antenna IBFD transceivers.

Regardless of the number of strong SI components, constructing the RF cancellation signal is essentially an interpolation problem since the delay of the SI component(s) is rarely precisely known. For this reason, the RF canceller should use several copies of the transmit signal with different delays to form an accurate replica of the observed SI **[99, 238]**, [28, 47]. The phases and/or amplitudes of these copies, or tap signals, are then adjusted, for instance, with VMs before combining them and forming the cancellation signal. Hence, the baseband-equivalent output signal of such an RF canceller can be written as follows:

$$y_{\mathrm{RFC}}(t) = y(t) - \sum_{i=0}^{N_{\mathrm{RFC}}-1} w_{i,\mathrm{RFC}}\, x_{\mathrm{PA}}(t-\tau_i), \tag{2.2}$$

where $w_{i,\mathrm{RFC}}$ is a complex coefficient denoting the phase and amplitude adjustment of the ith tap signal, N_{RFC} is the number of taps in the RF canceller, and τ_i is the fixed delay of the ith tap signal. The delay values are predetermined by measuring the impulse response of the coupling path and choosing the tap delays such that as much of the energy of the SI as possible is between the taps while minimizing also their mutual delays [47]. The correct phase and amplitude values can then be estimated adaptively by using, for instance, least mean squares (LMS)–type learning algorithms **[99, 238]**, [47]. Another RF cancellation approach is to only adjust the amplitudes of the tap signals, which means that the coefficients in $w_{i,\mathrm{RFC}}$ become real-valued. However, this requires more tap signals with a wider range of delays, thereby potentially resulting in a more complex implementation [28, 29].

In order to increase the flexibility of the RF canceller, Fig. 2.5b shows an alternative solution where the RF cancellation signal is in fact generated digitally, after which it is

upconverted to the RF domain with an auxiliary transmitter [63]. Since all the signal processing can be carried out in the digital domain, this allows for a more flexible and accurate generation of the RF cancellation signal, possibly incorporating also the effects of the multipath reflections into the modeling [20, 61, 63, 140]. However, as opposed to the architecture in Fig. 2.5a, the drawback of this solution is that the TX-induced impairments are not inherently included in the cancellation signal, meaning that they are not suppressed by the RF canceller unless explicitly modeled. Furthermore, an additional transmitter is also required, albeit without a PA [62]. The signal after this type of RF cancellation is as follows:

$$y_{\mathrm{RFC}}(t) = y(t) - \sum_{i=0}^{N_{\mathrm{RFC}}-1} w_{i,\mathrm{RFC}}\, x(t - iT_{\mathrm{s}}), \tag{2.3}$$

where $x(t)$ is the original transmit signal after the digital-to-analog converter (DAC), $w_{i,\mathrm{RFC}}$ is the complex coefficient of the ith tap signal, including also the amplification of the auxiliary receiver, and T_{s} is the sampling period used when generating the RF cancellation signal. This type of a solution has been widely studied in the earlier literature [22, 59, 61, 63, 140, 142], although in most cases the highest performance has so far been achieved with the RF cancellation structure shown in Fig. 2.5a **[238]**, [28]. Also note that the results included in this thesis have been generated using the RF cancellation structure of Fig. 2.5a, and hence PA output–based RF cancellation is assumed in all the derivations and system models in the continuation. Nevertheless, this does not decrease the applicability of the derived digital cancellation algorithms, as the chosen RF cancellation architecture only affects the relative power levels of the different signal components, not the signal models themselves **[P6]**.

Altogether, with sufficiently high RF cancellation performance, the SI is attenuated such that neither the LNA nor the ADC are saturated by the SI power, them being typically the critical components with regard to the highest tolerated power entering the RX chain **[P4]**. Having then amplified, downconverted, and digitized the total signal consisting of the residual SI and the signal of interest, digital cancellation is performed. Omitting the amplification occurring in the receiver for clarity, the signal after the digital canceller is expressed as follows:

$$y_{\mathrm{DC}}(n) = y_{\mathrm{ADC}}(n) - c(n) \approx s(n) + z(n), \tag{2.4}$$

where $y_{\mathrm{ADC}}(n)$ is the ADC output signal, $c(n)$ is the digital cancellation signal, and the index n is used to denote each sampling period. The digital cancellation signal is typically constructed from the original transmit data $x(n)$, using a predefined signal model and an estimate of the overall coupling channel. Ideally, the residual SI can be perfectly cancelled in the digital domain, after which only the received signal of interest, alongside with the noise, remains. The different digital cancellation solutions are described in detail in Chapter 4, where the possible structures for $c(n)$ are also disclosed.

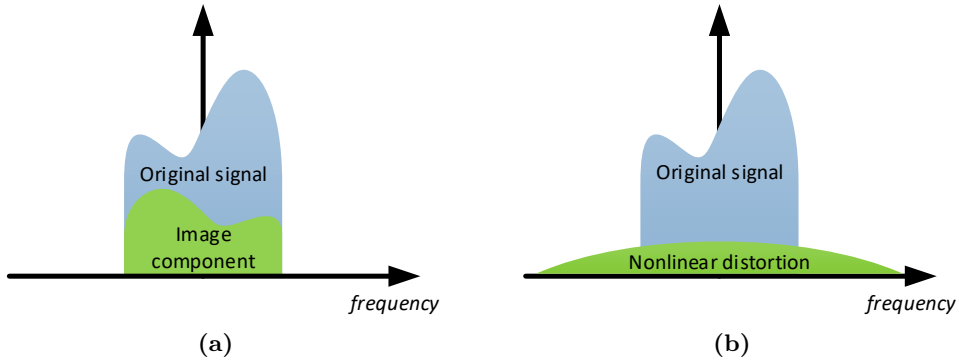


Figure 2.7: Frequency domain illustrations of (a) I/Q imbalance, and (b) nonlinear distortion.

2.4 Analog Imperfections

Due to the stringent cancellation requirements and the high power of the SI signal, many of the imperfections produced within the transceiver must be considered when modeling and regenerating the SI waveform for digital cancellation. This stems from the high power difference between the received SI signal and the signal of interest: even a mild distortion component in the former can be extremely powerful compared to the latter, which is weakened due to the much longer propagation distance. Hence, if such a distortion component is ignored in the cancellation processing, it will remain unaffected and therefore results in a heavily decreased SINR upon detection. Below, the most relevant analog imperfections are described and discussed.

2.4.1 I/Q Imbalance

I/Q imbalance is a prevalent issue in direct-conversion transceivers, stemming from the inherent phase and amplitude mismatches between the I- and Q-branches [39, 245]. As a result of these mismatches, a so-called image component is generated on top of the original signal, Fig. 2.7a illustrating this phenomenon in the frequency domain. The image component produced within a direct-conversion transceiver is in fact the original signal, whose spectrum has been inverted with respect to the frequency axis. In time domain, the image component is correspondingly the complex conjugate of the original signal. The magnitude of the image component in relation to the original signal is dictated by the severity of the I/Q imbalance, although typically it is clearly weaker than the original signal [69].

As opposed to the direct-conversion architecture, in the older *superheterodyne* transceivers, where an intermediate frequency is used during up- and downconversion, the mirror images are in fact in the adjacent frequency bands, and they can simply be filtered out either in the intermediate frequency (TX) or in the RF domain (RX). Consequently, the I/Q imbalance of a direct-conversion transceiver can be considered equivalent to the problem of image frequencies in the superheterodyne architecture. While the direct-conversion architecture avoids the need for the possibly bulky analog filters, it must

carefully match the phases and amplitudes of the I- and Q-branches to sufficiently suppress the image component.

Nevertheless, as already mentioned, the phases and amplitudes will inherently have some mismatches in all practical devices, and hence the image component should be considered when modeling the SI within a direct-conversion transceiver. In the transmitter, the signal at the I/Q mixer output can be expressed as follows:

$$x_{\mathrm{IQ}}^{\mathrm{TX}}(t) = k_{1,\mathrm{TX}}(t) \star x(t) + k_{2,\mathrm{TX}}(t) \star x^*(t), \tag{2.5}$$

where $k_{1,\mathrm{TX}}(t)$ is the response of the actual signal $x(t)$, and $k_{2,\mathrm{TX}}(t)$ is the response of the image component. Here, the term $k_{2,\mathrm{TX}}(t) \star x^*(t)$ represents the image component produced by the I/Q imbalance, and its magnitude is directly proportional to the severity of the phase and amplitude mismatches between the I- and Q-branches. To characterize and quantify the image suppression performance of a direct-conversion transmitter or receiver, the image rejection ratio (IRR) is typically used. It is defined as follows for the transmitter:

$$irr_{\mathrm{TX}}(f) = \frac{|K_{1,\mathrm{TX}}(f)|^2}{|K_{2,\mathrm{TX}}(f)|^2}, \tag{2.6}$$

where $K_{1,\mathrm{TX}}(f) = \mathcal{F}\{k_{1,\mathrm{TX}}(t)\}$ and $K_{2,\mathrm{TX}}(f) = \mathcal{F}\{k_{2,\mathrm{TX}}(t)\}$ are the frequency domain representations of the two responses. The I/Q imbalance occurring in the RX chain is modeled in an identical way. To give some insight into the magnitude of I/Q imbalance, the LTE UE specifications specify the lowest acceptable IRR as 25 dB [69]. While this is sufficient for HD systems, such I/Q imbalance levels result in a significant amount of residual SI in an IBFD transceiver if not properly addressed. This matter will be discussed in more detail in Chapter 3.

2.4.2 Nonlinear Distortion

Another fundamental issue, especially in low-cost communications devices, is the nonlinear distortion. It is primarily produced by the different active components, in particular the amplifiers, and can heavily distort the signal. In principle, it stems from some form of clipping, which results in the highest signal peaks being compressed. In other words, when driven sufficiently close to saturation, the gain of the amplifier is smaller when the amplitude of the input signal is higher, and thereby the relationship between the input and output signals is in fact nonlinear. In the frequency domain, nonlinear distortion can be illustrated as shown in Fig. 2.7b, where it exhibits itself as spectral regrowth. Note that, in the context of IBFD transceivers, only the nonlinear distortion falling onto the signal band needs to be considered since the distortion falling out-of-band can easily be filtered out in the receiver. The out-of-band nonlinearities are an important consideration only in terms of the transmitter spectral emission mask, which is typically defined in the system specifications to limit the interference produced on to the adjacent channels [69]. Nevertheless, since the focus of this work is only on the inband distortion, the spectral emission requirements are not explicitly considered.

In a typical case, the main source of nonlinear distortion is the TX PA **[P1, 18]**, [22, 144, 189]. The reason for this is the need for high power efficiency, while also having to

amplify the signal to the required transmit power level. These two requirements mean that the PA must operate close to its saturation point, which results in the nonlinear distortion of the waveform, especially with signals having high peak-to-average power ratio (PAPR) [105, 189]. Furthermore, even if the distortion is mild enough to fulfill the system specifications, it can still be extremely problematic for an IBFD device. For example, the strictest EVM requirement in the LTE specifications is 3.5% for the BS under the 256-QAM modulation scheme, which translates to a distortion component that is 29 dB weaker than the actual signal [68]. Making the reasonable assumption that the error component of a BS transmit signal is dominated by the PA nonlinearities [75, 76], this can be considered the highest allowed power level for the PA-induced nonlinear distortion under this modulation scheme. Hence, even with such a strict EVM requirement, the power level of the nonlinear distortion can easily be around 0 dBm at the TX output, meaning that in the RX chain it is likely orders of magnitude stronger than any signal of interest. What is more, when considering a lower order modulation scheme, the EVM requirements are less strict, meaning that the nonlinearities may be even stronger [68, 69]. This clearly indicates that nonlinear modeling of the TX PA is necessary in IBFD transceivers, as will be shown in more detail in Chapter 3.

In order to determine the power level of the nonlinear distortion more accurately, intercept points are commonly used. They describe the theoretical power level at which the nonlinear term of the considered order is equally powerful than the fundamental signal itself [79, p. 246]. This can then be used to calculate the power level of the nonlinear term for any reasonable component input power **[P1]**. As the 3rd-order nonlinearity is typically the dominant term, the 3rd-order input intercept point (IIP3) of the PA can be used to characterize the overall significance of the nonlinear distortion at the TX output. Then, the power level of the nonlinear distortion at the TX output can easily be approximated as follows **[P1]**:

$$p_{\mathrm{IMD,PA}} \approx \frac{p_{\mathrm{TX}}^3}{\left(iip3_{\mathrm{PA}}|k_{\mathrm{PA}}|^2\right)^2}, \tag{2.7}$$

where p_{TX} is the transmit power, $iip3_{\mathrm{PA}}$ is the IIP3 of the PA, and $|k_{\mathrm{PA}}|^2$ is the gain of the PA, all in linear units.

There are also various methods for modeling the PA-induced nonlinear distortion on a waveform level, such as the Volterra series or the Wiener model [1, 173, 241]. However, to limit the complexity of the model, and to ensure efficient parameter estimation, the widely-deployed parallel Hammerstein (PH) model is adopted in this thesis. In principle, a PH model refers to a system with parallel static nonlinearities, each having their own filter that models the memory effects [213]. In this thesis, the static nonlinearities are chosen to be monomials, while the memory effects are modeled using finite impulse response (FIR) filters. Denoting the baseband-equivalent PA input signal by $x_{\mathrm{PA}}^{\mathrm{in}}(n)$, its output signal can thereby be expressed with the adopted discrete-time PH model as follows **[P5, 15, 18]**, [57, 101, 241]:

$$x_{\mathrm{PA}}(n) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m=0}^{M_{\mathrm{PH}}} h_p(m) \left|x_{\mathrm{PA}}^{\mathrm{in}}(n-m)\right|^{p-1} x_{\mathrm{PA}}^{\mathrm{in}}(n-m), \tag{2.8}$$

where P is the nonlinearity order of the model, M_{PH} is the memory length of the model, and $h_p(m)$ contains the coefficients of the pth-order nonlinearity. Note that it is sufficient to include only the odd-order nonlinearities in the model when analyzing the inband distortion since the even-order terms fall outside the reception bandwidth [101]. This type of a model has been shown to be accurate for modeling a wide variety of practical PAs [15, 101, 173, 241], and in Chapter 5 it is shown to achieve high modeling accuracy also in the context of digital SI cancellation.

In addition to the nonlinear distortion produced in the transmitter, under some circumstances the RX chain can also distort the received signal in a nonlinear manner **[P1, 132]**, [10]. This typically occurs when the amount of SI suppression before the receiver is too low, resulting in the saturation of the RX LNA. It is possible to model and attenuate also the RX-induced nonlinearities **[132]**, [10], but the signal model becomes prohibitively complicated when also the PA is producing significant levels of distortion. Namely, then the overall coupling channel consists of the cascade of a nonlinearity, a wireless channel with memory, and another nonlinearity, resulting in an extremely large amount of nonlinear terms. For this reason, it is crucial to have sufficient RF cancellation performance since that will ensure that the power level of the receiver input signal is not high enough to produce significant nonlinear distortion in the LNA. This is also the underlying assumption in this thesis, and its validity is proven with the help of system calculations in Chapter 3 and further confirmed by the obtained measurement results reported in Chapter 5.

2.4.3 Analog-to-Digital Converter Quantization Noise

As opposed to the traditional HD systems, in IBFD transceivers also the receiver ADC plays a significant role [53, 204]. Namely, in a HD receiver, the accuracy of the quantization upon analog-to-digital conversion is rarely a bottleneck, assuming proper automatic gain control (AGC) that amplifies the overall signal to match the dynamic range of the ADC [79, p. 139]. There, the SINR is typically limited either by noise or by interference of some sort, as long as the ADC uses any reasonable amount of bits. However, the situation changes drastically when an IBFD transceiver is considered since then the ADC input signal also contains some residual SI. Even after a high amount of passive isolation and RF cancellation, the power of the residual SI can still be significantly higher than the power of the received signal of interest. This essentially means that the AGC is adjusting the gain based on the *residual SI*, not the signal of interest, to avoid clipping. Hence, upon the analog-to-digital conversion, the dynamic range of the signal of interest is well below the dynamic range of the ADC, meaning that it is quantized with less bits than in a corresponding HD device. Consequently, if the amount of analog SI suppression is not sufficiently high, the SINR of the signal of interest remains low due to the quantization effects, regardless of the digital cancellation performance [53, 204].

This phenomenon is illustrated in Fig. 2.8, where two scenarios are shown. In Fig. 2.8a, there is no residual SI at the ADC input, and consequently the whole dynamic range of the ADC can be used to quantize the signal of interest. As can be observed, the quantization effects are only minor and the signal quality remains good. On the other hand, in Fig. 2.8b the ADC input signal contains also some residual SI and, as a result,

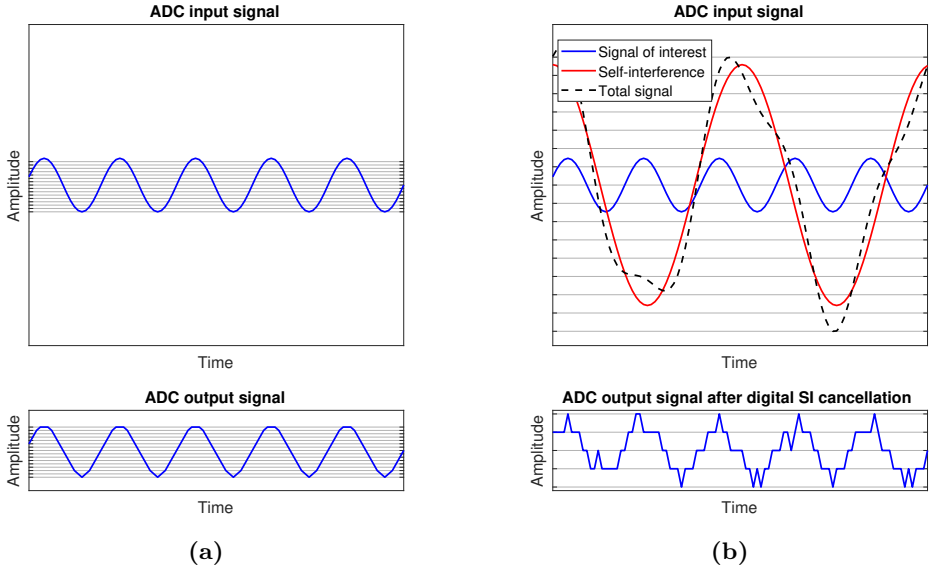


Figure 2.8: Effect of ADC quantization (a) without any SI, and (b) with a strong SI signal. The horizontal lines denote the quantization levels.

the AGC must use less gain in the receiver to avoid clipping in the ADC. Therefore, even after eliminating the SI with digital cancellation, the signal of interest is very noisy due to the quantization effects, as can be seen in the lower part of Fig. 2.8b. These examples show that, in the context of IBFD transceivers, the dynamic range of the ADC and the analog SI cancellation performance must be carefully considered to ensure sufficient SINR for the signal of interest in the digital domain.

The effect of the quantization noise can be determined by calculating its power level, which can be done if the voltage range of the ADC and the amount of bits it uses for quantizing the signal are known. Firstly, assuming an ideal AGC in the receiver that is capable of perfectly matching the dynamics of the input signal to the dynamic range of the ADC, the signal-to-quantization-noise ratio (SQNR) of an individual ADC can easily be shown to be as follows [38, p. 606]:

$$sqnr = \frac{3 \cdot 2^{2b}}{papr} = \frac{3 \cdot 2^{2b} p_{\mathrm{ADC}}}{p_{\mathrm{ADC}}^{\mathrm{max}}}, \tag{2.9}$$

where b is the number of bits used for the analog-to-digital conversion, $papr$ is the PAPR of the ADC input signal, $p_{\mathrm{ADC}}^{\mathrm{max}}$ is the peak power of the ADC input signal, and p_{ADC} is the average power of the ADC input signal, all in linear units. Considering then an I/Q receiver, which requires in fact *two* ADCs for digitizing both the I- and Q-signals, the overall quantization noise consists of the quantization noise produced by both ADCs. Using the expression in (2.9) to define the SQNRs of the I- and Q-branches as $sqnr_{\mathrm{I}}$ and $sqnr_{\mathrm{Q}}$, respectively, the absolute power level of the overall quantization noise in an I/Q

receiver can consequently be written as

$$p_{qn} = \frac{p_{ADC}^{I}}{sqnr_{\mathrm{I}}} + \frac{p_{ADC}^{Q}}{sqnr_{\mathrm{Q}}} = \frac{p_{ADC}^{max,I} + p_{ADC}^{max,Q}}{3 \cdot 2^{2b}} \approx \frac{V_{\mathrm{ADC}}^{2}}{3R2^{2b+1}} \approx \frac{p_{\mathrm{ADC}} papr}{3 \cdot 2^{2b}}, \tag{2.10}$$

where $p_{ADC}^{I}/p_{ADC}^{max,I}$ are the average and peak powers of the I-branch ADC input signal, respectively, $p_{ADC}^{Q}/p_{ADC}^{max,Q}$ are the corresponding quantities for the Q-branch, V_{ADC} is the peak-to-peak voltage range of an individual ADC, and R is the impedance of the system. Note that (2.10) assumes that the peak powers of the I- and Q-signals are roughly equal, and the last expression is based on the approximation that the peak power of the I/Q signal is roughly equal to the sum of the peak powers at the I- and Q-branches. These approximations facilitate the usage of only the I/Q signal–related quantities in evaluating the quantization noise power, and hence there is no need to consider the I- and Q-branches separately.

2.4.4 Transmitter Thermal Noise

Due to the extremely high power of the SI signal at the receiver input, even the transmitter thermal noise can result in an elevated interference floor in IBFD transceivers if not properly managed. In particular, the thermal noise present in the transmitter DAC output signal will also be amplified, alongside with the actual signal, and it is further magnified by the noise figure (NF) of the transmitter. This means that the TX-induced thermal noise can reach reasonably high power levels compared to the noise floor at the receiver input. In a HD transceiver, the transmitter noise can be neglected since the transmitter is either turned off during reception (cf. TDD), or it operates on a different frequency band and the noise can be filtered out (cf. FDD). However, in an IBFD transceiver, any noise included in the transmit signal will also overlap the received signal of interest in the frequency and time domains, meaning that it can potentially reduce the overall SINR. What is more, the transmitter thermal noise cannot obviously be modeled, which means that there are few options for canceling it in the receiver.

To ensure that the transmitter noise can be properly suppressed, the RF cancellation architecture illustrated in Fig. 2.5a is the preferable option. The reason for this is that in this type of an RF canceller the transmitter output signal is used as the cancellation signal, and hence it inherently includes also the transmitter noise. This means that the transmitter noise is attenuated by both the physical TX–RX isolation and the RF canceller, which together are typically enough to suppress it well below the receiver noise floor. On the other hand, if the auxiliary transmitter–based RF cancellation solution, illustrated in Fig. 2.5b, is used, the transmitter thermal noise realization is not included in the cancellation signal. This means that it is only attenuated by the passive isolation, which might not be sufficient to suppress it below the receiver noise floor. Hence, this speaks strongly for the transmitter output–based RF canceller as then also the transmitter noise is automatically taken care of. In Chapter 3, also the contribution of the transmitter thermal noise is quantified in more detail.

2.4.5 Oscillator Phase Noise

The effect of phase noise has also been widely studied in the context of IBFD transceivers [**233**], [163, 187, 195, 207]. Phase noise is caused by the varying phase of the LO signal during up- and downconversion, which results in a multiplicative distortion component. For the transmitter, the complex-valued baseband-equivalent phase noise model can be written as follows:

$$x_{\mathrm{PA}}^{\mathrm{in}}(t) = x_{\mathrm{IQ}}^{\mathrm{TX}}(t)e^{\jmath\phi_{\mathrm{tx}}(t)}, \tag{2.11}$$

where $x_{\mathrm{IQ}}^{\mathrm{TX}}(t)$ is the signal before upconversion, and $\phi_{\mathrm{tx}}(t)$ is the transmitter phase noise term. The latter is a stochastic process, whose statistics depend on the quality of the LO. In the ideal case where no phase noise is produced, $\phi_{\mathrm{tx}}(t)$ is merely a constant. In the receiver I/Q mixer, the signal is downconverted, which results in the following effective phase noise realization at the ADC input:

$$y_{\mathrm{ADC}}(t) = y_{\mathrm{IQ}}^{\mathrm{RX}}(t)e^{-\jmath\phi_{\mathrm{rx}}(t)}, \tag{2.12}$$

where $y_{\mathrm{IQ}}^{\mathrm{RX}}(t)$ is the signal before downconversion, and $\phi_{\mathrm{rx}}(t)$ is the random receiver phase noise term.

Under the assumption that $\phi_{\mathrm{tx}}(t)$ and $\phi_{\mathrm{rx}}(t)$ are independent of each other, phase noise is indeed a serious issue in IBFD transceivers, and it results in an elevated interference floor, even after all the cancellation stages [**233**], [163, 195, 207]. However, using independent LO signals in an IBFD transceiver is a rather pessimistic assumption since the transmitter and receiver operate on the same center frequency. Hence, in this case, the most sensible option is to use the same LO signal for both the up- and downconversion, and consequently $\phi_{\mathrm{tx}}(t) = \phi_{\mathrm{rx}}(t) = \phi(t)$.

Noting that the sign of the receiver phase noise term is opposite to that of the transmitter phase noise, it can be deduced that some of the phase noise is implicitly cancelled upon downconversion when using a shared TX/RX LO [**233**], [163]. For the self-cancellation of the phase noise in general, the deciding factor is the delay between the TX and RX I/Q mixers since that determines how well the phase noise during upconversion matches with the phase noise affecting the downconversion. Denoting the direct propagation delay between the TX and the RX I/Q mixers by τ_{PN}, the effective phase noise affecting the main SI component is as follows:

$$e^{\jmath\phi_{\mathrm{eff}}(t)} = e^{\jmath\phi(t)}e^{-\jmath\phi(t+\tau_{\mathrm{PN}})} = e^{\jmath(\phi(t)-\phi(t+\tau_{\mathrm{PN}}))}. \tag{2.13}$$

In particular, the common phase error (CPE), which refers to the mean value of $\phi(t)$, is perfectly cancelled upon downconversion since it can be expected to be static during the short propagation time. Also any frequency offset in the LO signal, commonly referred to as carrier frequency offset (CFO), is cancelled at this point since it is not affected by the delay. Hence, neither the CPE nor the CFO affect the overall SI waveform in an IBFD device with a shared LO, meaning that they can be omitted in this analysis.

Considering then the phase noise remaining after the self-cancellation, it can be further analyzed by making certain assumptions regarding the nature of the phase noise process. In particular, adopting the widely used free-running oscillator model where phase noise is modeled as a random-walk process, $\phi(t) = \sqrt{4\pi\beta_{3\mathrm{dB}}}B(t)$, where $B(t)$

denotes Brownian motion and $\beta_{3\text{dB}}$ is the 3-dB bandwidth of the phase noise [235, p. 16]. Hence, the effective phase noise can be expressed as follows:

$$\phi_{\text{eff}}(t) = \sqrt{4\pi\beta_{3\text{dB}}}\left(B(t) - B(t+\tau_{\text{PN}})\right). \tag{2.14}$$

Based on the basic properties of the Brownian motion, it can easily be shown that $\phi_{\text{eff}}(t) \sim \mathcal{N}\left(0, 4\pi\beta_{3\text{dB}}\tau_{\text{PN}}\right)$ [64, p. 301], where $\mathcal{N}\left(\mu, \sigma^2\right)$ denotes the normal distribution with mean μ and variance σ^2. It should also be noted that adopting this type of a model for the phase noise process can be considered a pessimistic scenario, as the phase noise performance of any real-world LO, typically utilizing some type of a phase-locked loop (PLL), is likely to be somewhat better [181, 244, 273].

In general, it can already be deduced that, with any reasonable delay, the effective phase noise of the main SI component is negligibly low [195]. It should be noted, however, that the self-cancellation of the phase noise is much weaker for the multipath SI components since their corresponding delays are longer. However, the multipath components themselves are also weaker due to the higher path loss, and hence the contribution of the effective phase noise remains negligible also in this case. These deductions are confirmed in Chapter 3, where it will be shown that, for a realistic propagation delay between the TX and RX I/Q mixers, the phase noise will have practically no effect on the overall residual SI, even when assuming such a pessimistic model for the phase noise process.

CHAPTER 3

Analysis of Analog Imperfections in Inband Full-Duplex Transceivers

In this chapter, the most significant analog imperfections in the context of IBFD transceivers are investigated and identified. The contents of this chapter are based on [**P1**–**P3**] as well as on the works in [**18, 82, 121, 123, 132, 133, 233**].

3.1 Power Levels of the Different Distortion Components

In order to determine the significance of the different impairments, an overall signal model for the considered IBFD transceiver architecture must first be derived. Here, the architecture of Fig. 2.5a with the PA output–based RF canceller is considered as it corresponds to the implemented prototype, discussed in more detail in Chapter 5, and is consequently the relevant scenario in terms of this thesis. The hereby obtained signal model can then be used to define the power levels of the different impairments. This reveals their relative significance within an IBFD transceiver and thereby facilitates the development of accurate SI cancellation algorithms. Moreover, since the objective is only to determine the power levels of the individual signal terms, a single-input and single-output (SISO) transceiver with frequency-independent models for the different components is assumed. However, the conclusions regarding the power levels of the different signal terms can also be applied as a baseline to MIMO systems and to frequency-dependent component models since these aspects do not have a major effect on the relative significance of the various impairments.

The system model used in calculating the power levels of the distortion components is illustrated in Fig. 3.1, where also the signals at the relevant parts of the transceiver chain are shown. For simplicity, the variable-gain amplifier (VGA) at the PA input is not explicitly considered in this analysis since it does not produce any significant nonlinear

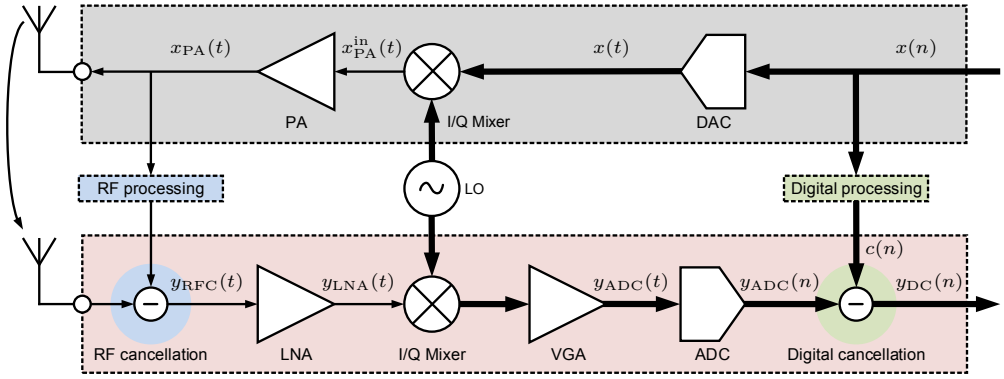


Figure 3.1: The system model used in determining the power levels of the different distortion components, alongside with the signals at the different stages of the IBFD transceiver.

distortion. Furthermore, all the variables are assumed to be in linear units, although their values are later specified in logarithmic units for clarity. Using the same notations as in Sections 2.3 and 2.4, the baseband-equivalent PA input signal can thereby be expressed as follows:

$$\begin{aligned} x_{\mathrm{PA}}^{\mathrm{in}}(t) &= \left(k_{1,\mathrm{TX}}\left[x(t)+z_{\mathrm{th}}(t)\right]+k_{2,\mathrm{TX}}\left[x^*(t)+z_{\mathrm{th}}^*(t)\right]\right)e^{\jmath\phi(t)} \\ &\approx \left[k_{1,\mathrm{TX}}x(t)+k_{2,\mathrm{TX}}x^*(t)\right]e^{\jmath\phi(t)}+k_{1,\mathrm{TX}}z_{\mathrm{th}}(t), \end{aligned} \tag{3.1}$$

where $k_{1,\mathrm{TX}}$ and $k_{2,\mathrm{TX}}$ are the total responses of the direct and image components before the PA, and $z_{\mathrm{th}}(t)$ is the thermal noise realization at the DAC output. Note that in this model $k_{1,\mathrm{TX}}$ and $k_{2,\mathrm{TX}}$ include the responses of all the components before the PA, such as that of the VGA. The approximation in (3.1) stems from ignoring the different distorted components of the DAC output noise, since they are negligibly weak in comparison to the other signal components. This signal is then fed to the PA, which distorts it in a nonlinear manner. The PA output signal is thereby as follows:

$$\begin{aligned} x_{\mathrm{PA}}(t) &= k_{\mathrm{PA}}x_{\mathrm{PA}}^{\mathrm{in}}(t)+x_{\mathrm{PA}}^{\mathrm{IMD}}(t)+z_{\mathrm{tx}}(t) \\ &\approx k_{\mathrm{PA}}\left[k_{1,\mathrm{TX}}x(t)+k_{2,\mathrm{TX}}x^*(t)\right]e^{\jmath\phi(t)}+x_{\mathrm{PA}}^{\mathrm{IMD}}(t)+k_{\mathrm{PA}}k_{1,\mathrm{TX}}\sqrt{F_{\mathrm{tx}}}z_{\mathrm{th}}(t), \end{aligned} \tag{3.2}$$

where k_{PA} is the frequency-independent response of the PA (i.e., its complex voltage gain), $x_{\mathrm{PA}}^{\mathrm{IMD}}(t)$ denotes the nonlinear distortion produced by the PA, $z_{\mathrm{tx}}(t)$ is the additional noise produced by the TX chain, and F_{tx} is the total noise factor[1] of the TX chain. Here, it is assumed that only the PA is producing additional noise in the transmitter, and hence its NF is equal to the NF of the whole transmitter. This assumption is done to simplify the analysis, since it does not affect the accuracy of the signal model in practice. Also note that, in reality, the NF of the TX chain does not amplify the existing thermal noise signal $z_{\mathrm{th}}(t)$; instead, the transmitter produces additional thermal noise such that the total noise power is increased by the NF. This simplification is done for both the TX and RX chains to clarify the notation as the interest is only in the power levels of the signal components.

[1]Note that the term *noise factor* refers to the linear quantity, while *noise figure (NF)* refers to the noise factor in logarithmic units.

The PA output signal is then attenuated either by the circulator, or by the path loss between the two antennas, after which it is fed to the RF canceller. Since the PA output signal is used for generating the RF cancellation signal, all the distortion components included in the transmit signal are also attenuated at this stage. After the RF canceller, the signal is amplified by the receiver LNA, which also produces some nonlinear distortion. The signal after the RF canceller and the RX LNA can thereby be written as follows:

$$\begin{aligned}
y_{\mathrm{LNA}}(t) &= k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} x_{\mathrm{PA}}(t) + y_{\mathrm{LNA}}^{\mathrm{IMD}}(t) + k_{\mathrm{LNA}} z'_{\mathrm{th}}(t) + z_{\mathrm{rx}}(t) \\
&\quad + k_{\mathrm{LNA}} z_{\mathrm{RFC}}(t) + k_{\mathrm{LNA}} s(t) \\
&\approx k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} k_{\mathrm{PA}} k_{1,\mathrm{TX}} x(t) e^{\jmath\phi(t)} + k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} k_{\mathrm{PA}} k_{2,\mathrm{TX}} x^*(t) e^{\jmath\phi(t)} \\
&\quad + k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} x_{\mathrm{PA}}^{\mathrm{IMD}}(t) + y_{\mathrm{LNA}}^{\mathrm{IMD}}(t) + k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} k_{\mathrm{PA}} k_{1,\mathrm{TX}} \sqrt{F_{\mathrm{tx}}} z_{\mathrm{th}}(t) \\
&\quad + k_{\mathrm{LNA}} \sqrt{F_{\mathrm{rx}}} z'_{\mathrm{th}}(t) + k_{\mathrm{LNA}} z_{\mathrm{RFC}}(t) + k_{\mathrm{LNA}} s(t),
\end{aligned} \tag{3.3}$$

where k_{LNA} is the complex gain of the RX LNA, a_{ANT} is the amount of physical isolation consisting either of the path loss between separate TX and RX antennas or the isolation provided by a circulator, a_{RFC} is the amount of RF cancellation, $y_{\mathrm{LNA}}^{\mathrm{IMD}}(t)$ is the nonlinear distortion produced by the RX LNA, $z'_{\mathrm{th}}(t)$ is the effective thermal noise realization in the RX chain, $z_{\mathrm{rx}}(t)$ is the additional noise produced in the receiver, $z_{\mathrm{RFC}}(t)$ is the noise produced by the RF canceller, and F_{rx} is the total noise factor of the RX chain. For notational clarity, it is assumed that the NF of the LNA is equal to the total NF of the receiver since the later components do not significantly contribute to the overall noise level.

After the LNA, the total received signal is downconverted and further amplified by the VGA to match its dynamics to the voltage range of the ADC. The input signal of the ADC can consequently be expressed as follows:

$$\begin{aligned}
y_{\mathrm{ADC}}(t) &= k_{\mathrm{VGA}} k_{1,\mathrm{RX}} y_{\mathrm{LNA}}(t) e^{-\jmath\phi(t+\tau_{\mathrm{PN}})} + k_{\mathrm{VGA}} k_{2,\mathrm{RX}} \left(y_{LNA}(t) e^{-\jmath\phi(t+\tau_{\mathrm{PN}})} \right)^* \\
&\quad + k_{\mathrm{VGA}} y_{\mathrm{MX}}^{\mathrm{IMD}}(t) + y_{\mathrm{VGA}}^{\mathrm{IMD}}(t) \\
&\approx k_{\mathrm{VGA}} k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} k_{\mathrm{PA}} k_{1,\mathrm{TX}} k_{1,\mathrm{RX}} x(t) e^{\jmath(\phi(t)-\phi(t+\tau_{\mathrm{PN}}))} \\
&\quad + k_{\mathrm{VGA}} \big(k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} k_{\mathrm{PA}} k_{2,\mathrm{TX}} k_{1,\mathrm{RX}} + k_{\mathrm{LNA}}^* a_{\mathrm{ANT}}^* a_{\mathrm{RFC}}^* k_{\mathrm{PA}}^* k_{1,\mathrm{TX}}^* k_{2,\mathrm{RX}} \big) x^*(t) \\
&\quad + k_{\mathrm{VGA}} k_{1,\mathrm{RX}} k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} x_{\mathrm{PA}}^{\mathrm{IMD}}(t) + k_{\mathrm{VGA}} k_{1,\mathrm{RX}} y_{\mathrm{LNA}}^{\mathrm{IMD}}(t) \\
&\quad + k_{\mathrm{VGA}} k_{1,\mathrm{RX}} k_{\mathrm{LNA}} a_{\mathrm{ANT}} a_{\mathrm{RFC}} k_{\mathrm{PA}} k_{1,\mathrm{TX}} \sqrt{F_{\mathrm{tx}}} z_{\mathrm{th}}(t) \\
&\quad + k_{\mathrm{VGA}} k_{1,\mathrm{RX}} k_{\mathrm{LNA}} \sqrt{F_{\mathrm{rx}}} z'_{\mathrm{th}}(t) + k_{\mathrm{VGA}} k_{1,\mathrm{RX}} k_{\mathrm{LNA}} z_{\mathrm{RFC}}(t) \\
&\quad + k_{\mathrm{VGA}} k_{1,\mathrm{RX}} k_{\mathrm{LNA}} s(t) + k_{\mathrm{VGA}} y_{\mathrm{MX}}^{\mathrm{IMD}}(t) + y_{\mathrm{VGA}}^{\mathrm{IMD}}(t),
\end{aligned} \tag{3.4}$$

where $k_{1,\mathrm{RX}}$ and $k_{2,\mathrm{RX}}$ are the frequency-independent responses of the direct and image components of the RX I/Q mixer, respectively, k_{VGA} is the complex gain of the VGA, $y_{\mathrm{MX}}^{\mathrm{IMD}}(t)$ is the nonlinear distortion produced by the RX I/Q mixer, and $y_{\mathrm{VGA}}^{\mathrm{IMD}}(t)$ is the nonlinear distortion produced by the RX VGA. Again, the I/Q images of the different distortion components have been ignored since they do not significantly contribute to the overall waveform. For the same reason, the phase noise affecting the distortion

components is also neglected since the phase noise of the linear SI term is dominant. Also, the linear SI term corresponding to the RX I/Q image of the TX I/Q image has been neglected since $|k_{2,\mathrm{TX}}k_{2,\mathrm{RX}}| \ll |k_{1,\mathrm{TX}}k_{1,\mathrm{RX}}|$ with any reasonable IRR [**P2**].

Finally, the ADC input signal is quantized, after which digital cancellation is performed. For the purposes of illustrating the magnitudes of the different impairment, it is assumed here that the digital canceller is only capable of suppressing the linear SI term, i.e., the one corresponding to $x(n)$. Thus, the cancellation signal is of the form $c(n) = \hat{h}^{\mathrm{L}}x(n)$, where $\hat{h}^{\mathrm{L}}$ denotes the one-tap estimate of the linear SI channel. Therefore, the signal fed to the receiver detector is as follows:

$$\begin{aligned}
y_{\mathrm{DC}}(n) &= y_{\mathrm{ADC}}(n) - c(n) = y_{\mathrm{ADC}}(n) - \hat{h}^{\mathrm{L}}x(n) \\
&\approx \left(k_{\mathrm{VGA}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}k_{1,\mathrm{RX}} - \hat{h}^{\mathrm{L}}\right)x(n) \\
&+ k_{\mathrm{VGA}}\left(k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{2,\mathrm{TX}}k_{1,\mathrm{RX}} + k^*_{\mathrm{LNA}}a^*_{\mathrm{ANT}}a^*_{\mathrm{RFC}}k^*_{\mathrm{PA}}k^*_{1,\mathrm{TX}}k_{2,\mathrm{RX}}\right)x^*(n) \\
&+ k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}x^{\mathrm{IMD}}_{\mathrm{PA}}(n) + k_{\mathrm{VGA}}k_{1,\mathrm{RX}}y^{\mathrm{IMD}}_{\mathrm{LNA}}(n) \\
&+ k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}\sqrt{F_{\mathrm{tx}}}z_{\mathrm{th}}(n) \\
&+ k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}\sqrt{F_{\mathrm{rx}}}z'_{\mathrm{th}}(n) + k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}z_{\mathrm{RFC}}(n) \\
&+ k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}s(n) + k_{\mathrm{VGA}}y^{\mathrm{IMD}}_{\mathrm{MX}}(n) + y^{\mathrm{IMD}}_{\mathrm{VGA}}(n) \\
&+ k_{\mathrm{VGA}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}k_{1,\mathrm{RX}}x(n)\jmath\phi_{\mathrm{eff}}(n) + z_{\mathrm{q}}(n),
\end{aligned} \tag{3.5}$$

where $z_{\mathrm{q}}(n)$ represents the quantization noise produced by the analog-to-digital conversion. Furthermore, it is assumed here that $e^{\jmath(\phi(t)-\phi(t+\tau_{\mathrm{PN}}))} = e^{\jmath\phi_{\mathrm{eff}}(n)} \approx 1 + \jmath\phi_{\mathrm{eff}}(n)$, which is a rather accurate approximation when $\phi_{\mathrm{eff}}(n)$ is small [235, p. 20]. This allows for the separation of the phase noise term from the linear SI term.

Using (3.5), the expressions for the power levels of the different signal/distortion components can then be determined, as is done in [**P1, P2**]. In particular, it is assumed that the complex gains of the different components, as well as the amounts of SI suppression in the different cancellation stages, are static and deterministic. Then, the power of the linear SI term can first be expressed as follows:

$$\begin{aligned}
p_{\mathrm{SI}} &= \mathbb{E}\left[\left|\left(k_{\mathrm{VGA}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}k_{1,\mathrm{RX}} - \hat{h}^{\mathrm{L}}\right)x(n)\right|^2\right] \\
&= \mathbb{E}\left[\left|\left(k_{\mathrm{VGA}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}k_{1,\mathrm{RX}} - \hat{h}^{\mathrm{L}}\right)\right|^2\right]p_{\mathrm{x}} \\
&= |a_{\mathrm{DC}}|^2\mathbb{E}\left[|k_{\mathrm{VGA}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}k_{1,\mathrm{RX}}|^2\right]p_{\mathrm{x}} \\
&= |k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2|a_{\mathrm{DC}}|^2p_{\mathrm{TX}},
\end{aligned} \tag{3.6}$$

where $p_x = \mathbb{E}[|x(n)|^2]$ is the power of the TX DAC output signal, $|a_{\mathrm{DC}}|^2$ is the amount of digital cancellation, and $p_{\mathrm{TX}} \approx |k_{\mathrm{PA}}|^2|k_{1,\mathrm{TX}}|^2p_{\mathrm{x}}$ is the transmit power of the device. The step from the second expression to the third expression follows directly from the definition for the amount of digital cancellation, a_{DC}, and its derivation is shown in [**P2**].

Correspondingly, the power of the I/Q imbalance–induced image component of the SI signal is as follows:

$$\begin{aligned} p_{\mathrm{SI,IM}} &= \mathbb{E}\Big[\big|k_{\mathrm{VGA}}\big(k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{2,\mathrm{TX}}k_{1,\mathrm{RX}} \\ &\qquad + k^*_{\mathrm{LNA}}a^*_{\mathrm{ANT}}a^*_{\mathrm{RFC}}k^*_{\mathrm{PA}}k^*_{1,\mathrm{TX}}k_{2,\mathrm{RX}}\big)\,x^*(n)\big|^2\Big] \\ &= |k_{\mathrm{VGA}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2|k_{\mathrm{PA}}|^2\left(|k_{2,\mathrm{TX}}|^2|k_{1,\mathrm{RX}}|^2 + |k_{1,\mathrm{TX}}|^2|k_{2,\mathrm{RX}}|^2\right)p_{\mathrm{x}} \\ &= |k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2\left(\frac{1}{irr_{\mathrm{TX}}}+\frac{1}{irr_{\mathrm{RX}}}\right)p_{\mathrm{TX}}, \end{aligned} \tag{3.7}$$

where irr_{TX} and irr_{RX} are the IRRs of the TX and RX chains, respectively. The latter expressions are obtained by assuming that the error of the RF cancellation signal is a circular random variable [**P2**], although the amount of obtained RF cancellation is still assumed to be static and deterministic.

Then, based on (2.7), and the discussion in Section 2.4.2 and [**P1**], the power of the PA-induced nonlinear distortion can be expressed as

$$\begin{aligned} p_{\mathrm{IMD,PA}} &= \mathbb{E}\left[\left|k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}x^{\mathrm{IMD}}_{\mathrm{PA}}(n)\right|^2\right] \\ &\approx \frac{|k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2p^3_{\mathrm{TX}}}{iip3^2_{\mathrm{PA}}|k_{\mathrm{PA}}|^4}, \end{aligned} \tag{3.8}$$

where $iip3_{\mathrm{PA}}$ is the IIP3 of the PA in linear power units. The final expression is simply obtained by assuming that the power of the PA-induced nonlinear distortion is dominated by the power of the 3rd-order intermodulation distortion, which means that its power can be approximated by using the IIP3 of the PA [**P1**].

The power of the RX-induced nonlinear distortion can be defined in a similar manner as follows [**P1**]:

$$\begin{aligned} p_{\mathrm{IMD,RX}} &= \mathbb{E}\left[\left|k_{\mathrm{VGA}}k_{1,\mathrm{RX}}y^{\mathrm{IMD}}_{\mathrm{LNA}}(n) + k_{\mathrm{VGA}}y^{\mathrm{IMD}}_{\mathrm{MX}}(n) + y^{\mathrm{IMD}}_{\mathrm{VGA}}(n)\right|^2\right] \\ &\approx |k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2\left(|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2p_{\mathrm{TX}}\right)^2\left[\left(\frac{|k_{\mathrm{LNA}}|^2}{iip2_{\mathrm{MX}}}+\frac{|k_{\mathrm{LNA}}|^2|k_{1,\mathrm{RX}}|^2}{iip2_{\mathrm{VGA}}}\right)\right. \\ &\quad \left.+\left(|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2p_{\mathrm{TX}}\right)\left(\frac{1}{iip3^2_{\mathrm{LNA}}}+\frac{|k_{\mathrm{LNA}}|^4}{iip3^2_{\mathrm{MX}}}+\frac{|k_{\mathrm{LNA}}|^4|k_{1,\mathrm{RX}}|^4}{iip3^2_{\mathrm{VGA}}}\right)\right], \end{aligned} \tag{3.9}$$

where $iip2_{\mathrm{MX}}$ and $iip2_{\mathrm{VGA}}$ are the 2nd-order input intercept points (IIP2s) of the RX I/Q mixer and VGA, respectively, and $iip3_{\mathrm{LNA}}$, $iip3_{\mathrm{MX}}$, and $iip3_{\mathrm{VGA}}$ are the IIP3s of the RX LNA, I/Q mixer, and VGA, respectively. Upon deriving the power level of the RX-induced nonlinearities, it has again been assumed that it is sufficient to consider only 2nd- and 3rd-order distortion to accurately determine the total power. Furthermore, only the I/Q mixer and the VGA produce 2nd-order nonlinear distortion that falls onto the RX frequency band [**P1**].

The power of the TX-induced thermal noise, including also the noise produced by the RF canceller, can then be expressed as follows:

$$\begin{aligned} p_{\mathrm{n,TX}} &= \mathbb{E}\left[\left|k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}\left(a_{\mathrm{ANT}}a_{\mathrm{RFC}}\sqrt{F_{\mathrm{tx}}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}z_{\mathrm{th}}(n)+z_{\mathrm{RFC}}(n)\right)\right|^2\right] \\ &= |k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2\left(F_{\mathrm{tx}}|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2|k_{\mathrm{PA}}|^2|k_{1,\mathrm{TX}}|^2p_{\mathrm{th}}+p_{\mathrm{RFC}}\right), \end{aligned} \quad (3.10)$$

where p_{th} is the thermal noise power over the considered bandwidth, and p_{RFC} is the noise power at the RF canceller output. The thermal noise power in room temperature and dBm units is given by $p_{\mathrm{th}}^{\mathrm{dBm}} = -174 + 10\log_{10}(W)$, where W is the bandwidth of the system **[P1]**. Correspondingly, the power level of the RX-induced thermal noise is written as follows:

$$\begin{aligned} p_{\mathrm{n,RX}} &= \mathbb{E}\left[\left|k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}\sqrt{F_{\mathrm{rx}}}z'_{\mathrm{th}}(n)\right|^2\right] \\ &= F_{\mathrm{rx}}|k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2p_{\mathrm{th}}. \end{aligned} \quad (3.11)$$

The power of the phase noise–induced SI term is then expressed as

$$\begin{aligned} p_{\mathrm{PN}} &= \mathbb{E}\left[|k_{\mathrm{VGA}}k_{\mathrm{LNA}}a_{\mathrm{ANT}}a_{\mathrm{RFC}}k_{\mathrm{PA}}k_{1,\mathrm{TX}}k_{1,\mathrm{RX}}x(n)\jmath\phi_{\mathrm{eff}}(n)|^2\right] \\ &= |k_{\mathrm{VGA}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2|k_{\mathrm{PA}}|^2|k_{1,\mathrm{TX}}|^2|k_{1,\mathrm{RX}}|^2\mathbb{E}[|\phi_{\mathrm{eff}}(n)|^2]p_{\mathrm{x}} \\ &= 4\pi|k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2\beta_{\mathrm{3dB}}\tau_{\mathrm{PN}}p_{\mathrm{TX}}, \end{aligned} \quad (3.12)$$

where the final expression is based on the earlier observation that $\phi_{\mathrm{eff}}(t)\sim\mathcal{N}(0,4\pi\beta_{\mathrm{3dB}}\tau_{\mathrm{PN}})$. Thus, (3.12) can be used to evaluate the power of the phase noise–induced SI term, given the delay between the TX and RX mixers and the 3-dB bandwidth of the phase noise process.

Then, using (2.10), the power of the quantization noise can be written as follows:

$$\begin{aligned} p_{\mathrm{QN}} &= \mathbb{E}\left[|z_{\mathrm{q}}(n)|^2\right] = \frac{V_{\mathrm{ADC}}^2}{3R\,2^{2b+1}} = \frac{p_{\mathrm{ADC}}papr}{3\cdot 2^{2b}} \approx \frac{\frac{p_{\mathrm{SI}}}{|a_{\mathrm{DC}}|^2}papr}{3\cdot 2^{2b}} \\ &= \frac{|k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2paprp_{\mathrm{TX}}}{3\cdot 2^{2b}}. \end{aligned} \quad (3.13)$$

The last two expressions stem from the assumptions that *(i)* the linear SI term is dominating the power of the total signal upon the analog-to-digital conversion and *(ii)* the AGC has perfect knowledge of the necessary signal powers so that the full voltage range of the ADC is utilized, both of which are reasonable assumptions for the purposes of this analysis. The gain of the VGA follows then directly from the expressions in (3.13). In particular, solving for $|k_{\mathrm{VGA}}|^2$ in (3.13), we get

$$\begin{aligned} |k_{\mathrm{VGA}}|^2 &= \frac{p_{\mathrm{ADC}}}{|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2p_{\mathrm{TX}}} \\ &= \frac{V_{\mathrm{ADC}}^2}{2R\,|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2|a_{\mathrm{ANT}}|^2|a_{\mathrm{RFC}}|^2paprp_{\mathrm{TX}}}. \end{aligned} \quad (3.14)$$

Finally, the power of the received signal of interest is expressed as

$$p_{\mathrm{SOI}} = \mathbb{E}\left[|k_{\mathrm{VGA}}k_{1,\mathrm{RX}}k_{\mathrm{LNA}}s(n)|^2\right] = |k_{\mathrm{VGA}}|^2|k_{1,\mathrm{RX}}|^2|k_{\mathrm{LNA}}|^2p_{\mathrm{RX}}, \quad (3.15)$$

where $p_{\mathrm{RX}} = \mathbb{E}[|s(n)|^2]$ is the received signal power at the input of the RX chain.

Table 3.1: The essential default system parameters used in evaluating the different power levels, unless otherwise mentioned.

Parameter	**Value**
System bandwidth (W)	20 MHz
Power of the received signal of interest (p_{RX})	-85 dBm
PAPR of the transmit signal ($papr$)	10 dB
Amount of antenna isolation ($\lvert a_{\mathrm{ANT}}\rvert^2$)	-40 dB
Amount of RF cancellation ($\lvert a_{\mathrm{RFC}}\rvert^2$)	-35 dB
RF canceller output noise, incl. coupling loss (p_{RFC})	-97 dBm
Number of bits in the RX ADCs (b)	12
Peak-to-peak voltage range of the RX ADCs (V_{ADC})	4 V
Impedance of the system (R)	50 Ω
TX DAC output power (p_{x})	-20 dBm
TX & RX IRR (irr_{TX} / irr_{RX})	25 dB
Transmit power (p_{TX})	23 dBm
3-dB bandwidth of the phase noise (β_{3dB})	50 Hz
Delay between TX and RX I/Q mixers (τ_{PN})	3 ns

Table 3.2: The essential default component parameters used in evaluating the different power levels, unless otherwise mentioned.

Component		**Gain [dB]**	**IIP2 [dBm]**	**IIP3 [dBm]**	**NF [dB]**
TX	I/Q Mixer & VGA	0–21 ($\lvert k_{1,\mathrm{TX}}\rvert^2$)	–	–	10
TX	PA	24 ($\lvert k_{\mathrm{PA}}\rvert^2$)	–	20 ($iip3_{\mathrm{PA}}$)	5
TX	Total	24–45	–	-1	10.0 (F_{tx})
RX	LNA	25 ($\lvert k_{\mathrm{LNA}}\rvert^2$)	–	-9 ($iip3_{\mathrm{LNA}}$)	4.1
RX	I/Q Mixer	6 ($\lvert k_{1,\mathrm{RX}}\rvert^2$)	42 ($iip2_{\mathrm{MX}}$)	15 ($iip3_{\mathrm{MX}}$)	4
RX	VGA	0–69 ($\lvert k_{\mathrm{VGA}}\rvert^2$)	43 ($iip2_{\mathrm{VGA}}$)	14 ($iip3_{\mathrm{VGA}}$)	4
RX	Total	31–100	11	-17	4.1 (F_{rx})

3.2 Evaluating the Distortion Power Levels with Realistic System Parameters

The power levels of the different signal and distortion components are next evaluated using (3.6)–(3.15) and some example system parameters. Here, we adopt similar parameters as in **[P1]**, taken mostly from [79, 191] and listed in Tables 3.1 and 3.2. These parameter values have been used in generating the forthcoming figures, unless otherwise mentioned. Moreover, for the purposes of determining the most significant RF impairments, the linear digital canceller is assumed to be capable of perfectly suppressing the linear SI term, meaning that $a_{\mathrm{DC}} = 0$, and consequently also $p_{\mathrm{SI}} = 0$.

Also, as already discussed when deriving the equations for the power levels, only the total NFs of the TX and RX chains are considered in the modeling, the NFs of the individual components being used for calculating the total NFs. Furthermore, the output noise of the RF canceller is determined based on the output noise level of a particular VM [89], which is used in the prototype implementation presented in [**P4**]. The RF cancellation signal is also assumed to be combined with the received signal via a 10-dB directional coupler, incurring a 10-dB decrease in the VM output noise level.

The power of the phase noise is calculated by assuming a certain delay between the TX and RX LOs, which defines the amount of self-cancellation. In this thesis, a delay of 3 ns, corresponding to a propagation distance of roughly 90 cm, is assumed, which likely represents a pessimistic estimate of the delay between the TX and RX I/Q mixers. The 3-dB bandwidth of the phase noise process is chosen based on the value used in [**233**] and, together with the delay, it defines the magnitude of the residual phase noise. It should be noted that, although the self-cancellation effect is significantly weaker for the multipath components, they are much weaker to begin with and hence the phase noise corresponding to these multipath reflections is omitted in this analysis.

Firstly, Fig. 3.2 shows the power levels of the signal components with respect to the transmit power. It can be observed that the SINR is compromised even with the lowest considered transmit power of 5 dBm, since the power of the SI image component is already then higher than the noise floor. Furthermore, when the transmit power is increased, the significance of the image component gets even higher. Hence, even though the IRR is fulfilling the LTE specifications [69], I/Q imbalance is still the most significant source of distortion in an IBFD device, and must consequently be incorporated into the SI modeling [**P2**]. Moreover, with the higher transmit powers, also the PA-induced nonlinear distortion becomes a considerable factor. Especially, with the highest considered transmit power of 25 dBm, it is already heavily decreasing the SINR. This indicates that also the transmitter nonlinearities should be modeled and cancelled in an IBFD device [**P3**–**P5, 125**], [10, 28]. Also note that, in Fig. 3.2, some of the power levels in fact decrease as the transmit power is increased. This is due to the AGC, which must lower the gain with the higher transmit powers to avoid clipping in the ADC as the linear SI power is then obviously stronger.

Another important observation from Fig. 3.2 is the fact that, with the considered example system parameters, the TX-induced thermal noise is in fact at the same level as the receiver noise floor. This means that the effective overall receiver noise floor of the considered IBFD device is roughly 3 dB higher than that of the corresponding HD receiver. This noise component is dominated by the RF canceller output noise, and consequently it can be largely eliminated by ensuring that the RF canceller is producing sufficiently little noise. Hence, this is also an important aspect to consider when designing an IBFD transceiver. Nevertheless, it should also be noted that a 3-dB decrease in the SINR is in general not intolerable for an IBFD transceiver since it can still provide a throughput gain under such circumstances. This stems from the fact that the SINR affects the capacity inside a logarithm, while the twofold gain of the IBFD operation is outside the logarithm, as discussed already in Chapter 1. As for the other sources of impairments, such as phase noise, RX-induced nonlinearities, and quantization noise, Fig. 3.2 indicates that they do not significantly contribute to the overall distortion power with these system parameters.

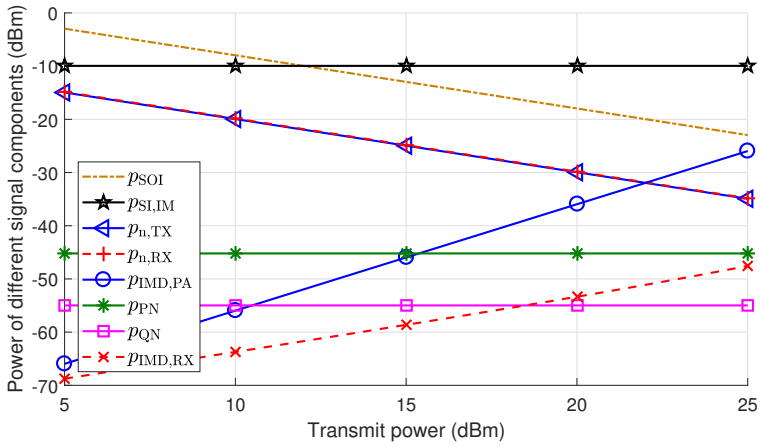


Figure 3.2: Power levels of the different signal components with respect to the transmit power.

To gain further insight into the effect of the-PA induced nonlinearities, let us define the SINR after digital cancellation as

$$sinr_{\mathrm{DC}} = \frac{p_{\mathrm{SOI}}}{p_{\mathrm{SI}} + p_{\mathrm{SI,IM}} + p_{\mathrm{IMD,PA}} + p_{\mathrm{IMD,RX}} + p_{\mathrm{n,TX}} + p_{\mathrm{n,RX}} + p_{\mathrm{PN}} + p_{\mathrm{QN}}}, \quad (3.16)$$

which assumes for simplicity that all the distortion components are uncorrelated. Assuming then that also the SI image component can be perfectly cancelled in the digital domain, in addition to the linear SI component, Fig. 3.3 shows the resulting SINR where the nonlinear distortion produced by the PA is now the dominant impairment. The SINR has been plotted there with respect to the IIP3 of the PA, using five different RF cancellation levels. It can be observed that a reasonably high IIP3 is required to prevent excessive SINR loss in an IBFD transceiver, especially when the amount of RF cancellation is 40 dB or less. Namely, with 40 dB of RF cancellation, the PA IIP3 must be at least 25 dBm to ensure that the PA-induced nonlinearities are not limiting the receiver performance. If the amount of RF cancellation is 30 dB, the IIP3 must be close to 30 dBm to ensure no SINR loss due to the PA nonlinearities, although then also other distortion components start to limit the SINR.

In general, these findings indicate that being capable of suppressing the PA-induced nonlinear distortion in the digital domain is greatly beneficial, as many of the lower-cost PAs have IIP3s less than 15 dBm [166, 223, 242]. Based on Fig. 3.3, this is insufficient for guaranteeing the linearity of the observed SI signal under reasonable RF cancellation performance, resulting in a decreased SINR when the nonlinearity is not modeled. Hence, the nonlinear distortion of the SI waveform should be incorporated into the signal models used for digital cancellation [**P1**].

Then, to demonstrate the effect of phase noise on the SINR, Fig. 3.4 shows the SINR loss caused by the phase noise with respect to the transmit power. In particular, the SINR is calculated according to (3.16) using four different 3-dB bandwidths for the phase noise, and then compared to a scenario where $\beta_{\mathrm{3dB}} = 0$ Hz, i.e., $p_{\mathrm{PN}} = 0$. Again, it is assumed that both the linear SI term and the SI image component can be

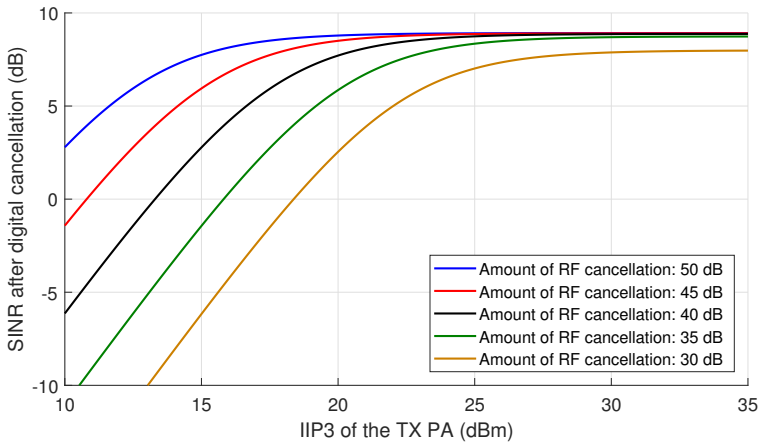


Figure 3.3: SINR after digital cancellation with respect to the IIP3 of the TX PA, shown for different RF cancellation levels. It is assumed here that both the linear SI and the SI image component can be perfectly cancelled in the digital domain.

perfectly cancelled in the digital domain, while the other terms remain unaffected in the final cancellation stage. Moreover, for reference, the SINR loss is also shown for a fully linear PA. It can firstly be observed that, for the nonlinear PA, the effect of phase noise is indeed relatively small, even with a very large 3-dB bandwidth. Namely, with $\beta_{3\text{dB}} = 300$ Hz, the SINR loss remains below 0.5 dB for all the considered transmit powers, while it is still only a little bit more than a decibel at worst with the highest considered 3-dB bandwidth of 1 kHz. Considering that a free-running oscillator is assumed when deriving the power of the phase noise–induced SI term, the significance of phase noise is likely to be even lower for a more realistic model.

The peculiar form of the SINR loss curves for the nonlinear PA is explained by the PA-induced nonlinearities, which become more and more dominant with higher transmit powers. That is, with the lower transmit power values, the phase noise levels have a larger impact on the SINR loss, while the PA nonlinearities become stronger when the transmit power is increased. This results in the residual SI power being dominated by the nonlinear distortion with the highest transmit powers, meaning that the contribution of the phase noise to the SINR loss becomes less and less significant. Consequently, the SINR loss due to phase noise converges to 0 dB when the transmit power is increased sufficiently high, while the SINR loss with a linear PA keeps on increasing since the phase noise remains the dominant distortion component.

To conclude, the findings regarding the power levels of the different SI terms clearly indicate that both the PA-induced nonlinearities and the I/Q imbalance–induced SI image component need to be considered in an IBFD device **[P1, P2]**. The other analog impairments do not significantly contribute to the overall SI power, although they must still be considered to some extent when designing and dimensioning IBFD transceivers. Namely, the ADCs must have a sufficient number of bits while the LO must be of sufficiently high quality, or else the SI modeling accuracy and the overall signal quality might be affected by these impairments. Moreover, also the thermal noise must be taken

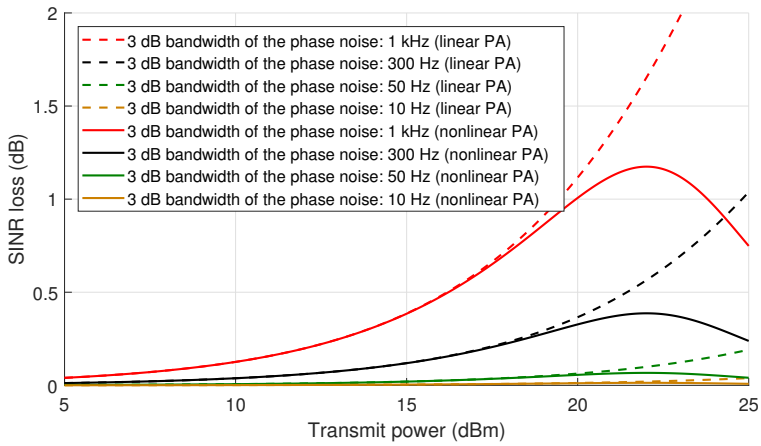


Figure 3.4: The phase noise–induced SINR loss with respect to the transmit power, shown for different 3-dB bandwidths. It is assumed here that both the linear SI and the SI image component can be perfectly cancelled in the digital domain.

into account, especially when designing the RF canceller. For the above reasons, this thesis presents different signal models that incorporate the PA-induced nonlinearities and/or the I/Q imbalance, and which can then be used for efficient digital SI cancellation. They are derived in Chapter 4 and evaluated in Chapter 5.

CHAPTER 4

Digital Self-interference Cancellation under Analog Imperfections

This chapter presents different signal models to be used for digital SI cancellation, alongside with two alternative parameter learning algorithms for estimating the necessary SI channel coefficients. The objective of these signal models is to allow for accurate reconstruction of the SI signal observed in the digital domain under analog imperfections, thereby facilitating high SI cancellation performance. The contents of this chapter are based on the journal publications in [**P2**–**P6**], as well as on the works in [**4, 17, 18, 120**–**128, 133**].

4.1 Background and State of the Art

Modeling of the residual SI signal in the digital domain is a crucial aspect for an IBFD transceiver as the objective of the digital canceller is typically to suppress the SI signal below the receiver noise floor. Furthermore, as shown in Chapter 3, in many cases this requires the modeling of some of the RF impairments since otherwise the accuracy of the cancellation signal is not sufficiently high. Nevertheless, in many of the related works, a linear signal model has been assumed in the digital cancellation stage [13, 44, 50, 51, 63, 240, 256], and consequently none of the RF impairments have been modeled. This represents a baseline for the signal model used within a digital canceller, and it is also described in detail in Section 4.2.1 below. However, as opposed to some works where the SI is cancelled in the frequency domain [13, 50, 51], in this thesis only time-domain cancellation is considered.

To improve the accuracy, many of the reported digital cancellation solutions incorporate also a model for the nonlinear TX PA [**P3**–**P5, 17, 18, 124, 125**], [10, 22, 28, 37, 66, 156, 219]. Considering that in most systems the PA-induced nonlinearities are indeed the dominant source of distortion, such a nonlinear digital canceller is typically capable of highly efficient SI cancellation [**P4, P5**], [28]. In the aforementioned works, a PH

model is primarily used as it incorporates also the different memory effects, as described in Section 2.4.2. Furthermore, the works in [22, 219] utilize the nonlinear signal model for predistorting the PA output signal, which means that a linear model can be used at the actual digital cancellation stage. Addressing the PA-induced nonlinear distortion by intentionally introducing a polarization mismatch between the TX output and the RX digital domain has also been considered [269]. Moreover, in [37], also the nonlinear distortion produced by the RX chain is incorporated into the overall digital cancellation signal model, in addition to the TX nonlinearities, albeit only 3rd-order distortion is considered for simplicity.

In addition to the PA nonlinearities, also the I/Q imbalance can be an issue in a low-cost radio transceiver, as observed in Chapter 3. To this end, [34] investigates the performance of spatial-domain suppression techniques under TX/RX nonlinearities and I/Q imbalance, while the work in [32] proposes a time-domain digital cancellation solution that incorporates these impairments into the signal model. Moreover, a transmit beamforming solution capable of modeling also the I/Q imbalance is proposed in [96], whereas [209] presents a widely linear signal model for the SI that includes the effect of TX I/Q imbalance. The significance of I/Q imbalance is also observed in [26], where a time-domain digital canceller modeling both I/Q imbalance and DAC nonlinearities is proposed. Using a measured SI signal, it is shown to improve the amount of digital cancellation by 10 dB compared to other state-of-the-art solutions. A digital SI canceller capable of modeling nonlinear distortion and I/Q imbalance is also proposed in [243], while **[17]** presents a signal model for the residual SI observed in a MIMO IBFD transceiver that incorporates both the PA-induced nonlinear distortion and the TX/RX I/Q imbalance. The joint model in **[17]** is essentially a special case of the signal model presented in **[P6]** (and Section 4.2.4 below) as it neglects the crosstalk between the transmitters, whereas the model derived in **[P6]** incorporates also the crosstalk effects. In addition to these, SI signal models that consider both PA nonlinearity and I/Q imbalance are presented also in [162, 164, 179], although these works only assume a 3rd-order model for the PA. However, in [179] the phase noise effects are also compensated for, in addition to the PA-induced nonlinearities and the I/Q imbalance.

In fact, there are also various other studies where the impact of phase noise is analyzed in the context of digital SI cancellation **[233]**, [7, 11, 74, 143, 154, 163, 195, 206, 207, 216, 234]. A typical assumption among these works is that the TX and RX chains employ separate LOs with independent phase noise characteristics, and hence no self-cancellation occurs [7, 11, 74, 143, 154, 206, 207, 234]. However, as observed in Chapter 3, under the sensible assumption of a shared TX/RX LO, and when considering a realistic delay between the up- and downconversion stages, phase noise does not significantly contribute to the overall SI waveform. Consequently, it can be neglected when deriving a signal model for the SI in the digital domain without incurring a significant reduction in the modeling accuracy.

In addition to these impairments, also the nonlinear distortion produced within the RX chain has been considered in some of the works **[133]**, [10]. Modeling the nonlinearity at this stage is somewhat more challenging than in the transmitter since the input signal of the nonlinearity is not precisely known due to the unknown SI coupling channel. This results in either an extremely complicated linear-in-parameters signal model, or in a rather involved two-stage estimation procedure **[133]**. Considering the findings of

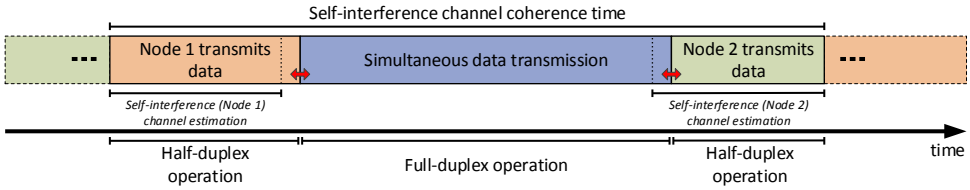


Figure 4.1: The frame structure of a flexible communications procedure for a bidirectional link between two IBFD nodes, studied in [**128**].

Chapter 3 which suggest that, with reasonable RF cancellation performance, the input power of the RX chain is usually sufficiently low for the RX-induced nonlinearities to be negligibly weak, they are not included in the cancellation signal models presented in this thesis.

Furthermore, there are also studies that propose using an auxiliary RX chain for obtaining a reference signal for the digital canceller from the TX output, since this allows the usage of *linear processing* while still being capable of canceling the TX-induced impairments [**123**], [6, 74, 144, 145, 267]. The benefit of this type of an architecture is the lower computational cost of digital SI cancellation in general, as none of the TX impairments need to be explicitly considered. By incorporating a model for the RX-induced nonlinearities, such an architecture is shown to obtain good SI cancellation performance [6]. However, an additional RX chain is obviously required for downconverting and digitizing the TX output signal, which means that the amount of required RF hardware is larger. Hence, because of this drawback, implementing a digital canceller where the different impairments are explicitly modeled is in many respects more intriguing.

As for learning the parameters of the signal model utilized for digital SI cancellation, a significant aspect is whether the estimation should be done during a dedicated training period when there are no other signals being received, or if the parameters can be learned while also receiving a signal of interest [**122**], [36, 144, 164]. Although the derivations and results within this thesis are done under the assumption that the device receives only its own SI, this aspect has been studied extensively in [**122, 127, 128**]. Especially, [**128**] considers a scenario where two IBFD capable devices communicate bidirectionally, taking also into account the effect of SI channel estimation. The analyzed communications procedure is illustrated in Fig. 4.1, where the two IBFD nodes exchange data both in HD and IBFD modes, the SI channel estimation being performed at least partially during the former. A key insight is that the signals transmitted during the HD or training period can also contain useful data, as the IBFD device obviously knows anyway what it is transmitting. Consequently, these HD calibration periods, during which the SI channel estimation occurs, also contribute to the overall data rate of the system. The data rate regions of the considered two-node system are determined in [**128**] by adjusting the lengths of the two HD periods and then evaluating the resulting SINRs with realistic waveform simulations that incorporate also the effects of all the relevant RF impairments. The findings indicate that the largest rate region is typically obtained when the SI parameter estimation can be performed at least partially during a HD period. A similar analysis is carried out in [**127**] for an IBFD relay, where the same

conclusions are reached. Hence, these works indicate that having a dedicated calibration period is typically the best option for learning the parameters of the SI signal model.

A somewhat different approach into the SI channel estimation is taken, for instance, in [143, 144, 163, 164], where the received signal of interest is in fact incorporated into the overall signal model for digital cancellation, instead of considering it merely as noise from the perspective of SI parameter estimation. Using then a semi-blind and iterative estimation procedure, channel coefficients for both the SI and the received signal of interest can be obtained. Although such estimation procedures are shown to perform relatively well, especially in the presence of a strong received signal of interest [163], in this thesis a scenario with a separate calibration period is considered. This is a justifiable approach based on the findings of **[127, 128]**, which suggest that using such HD periods for SI parameter estimation is the most viable option when the signal of interest is considered as noise in the SI signal model. Extending the proposed cancellation and parameter estimation solutions to also explicitly incorporate the received signal of interest into the SI modeling is left as a future work item.

All in all, the related literature contains various works where also the RF impairments have been considered in the digital cancellation stage. This further demonstrates that modeling of the RF impairments is indeed an important aspect in IBFD transceivers. In the following, different signal models for the residual SI in the digital domain are presented, after which the parameter estimation–related aspects are discussed. Together, these topics will provide a complete and thorough description of a digital SI canceller capable of modeling various different RF impairments.

4.2 Advanced Signal Models for Self-interference

In this thesis, three different advanced SI signal models are presented and used for digital cancellation, in addition to the basic linear signal model which is used as a benchmark. Furthermore, all the signal models are derived for a MIMO IBFD transceiver, which means that they can be readily applied to a multi-antenna system. However, the measurement-based evaluation, presented in Chapter 5, is done with a SISO transceiver since that is already sufficient to show the accuracy of the different signal models. Also note that, as mentioned above, in this chapter it is assumed that only the SI is being received, i.e., there is no signal of interest present in the total RX signal. Nevertheless, the derivations are also valid for a case without a calibration period as then the signal of interest can be considered a part of the overall noise.

4.2.1 Linear Signal Model

The most basic approach in modeling the residual SI in the digital domain is to assume a perfectly linear transceiver chain. That is, all the impairments are neglected and only the different memory effects, produced by the transceiver and the wireless coupling channel, are considered. Such a linear system can be modeled as illustrated in Fig. 4.2. In this case, the baseband-equivalent output signal of the jth transmitter can be written as follows:

$$x_{j,\mathrm{PA}}(t) = f_{j,\mathrm{TX}}(t) \star x_j(t) + e_{j,\mathrm{TX}}(t), \tag{4.1}$$

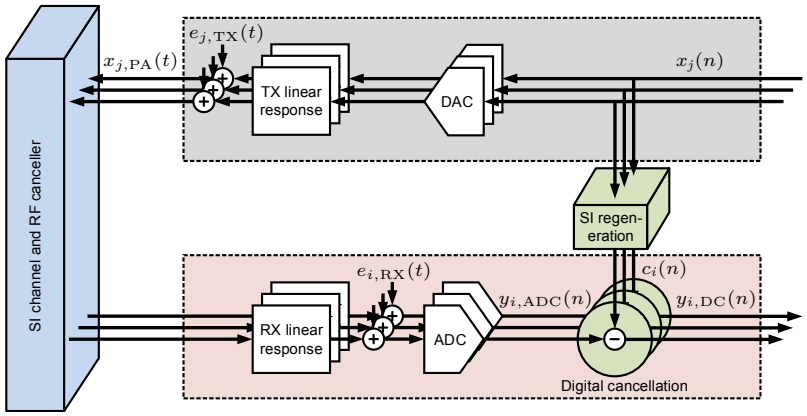


Figure 4.2: The system model used in deriving the linear signal model for the digital canceller.

where $f_{j,\mathrm{TX}}(t)$ denotes the general frequency-dependent response of the jth transmitter, including the linear gain and the possible memory effects, and $e_{j,\mathrm{TX}}(t)$ represents all the unmodeled distortion components and the transmitter noise, i.e., the modeling error.

Then, the signals are transmitted and consequently received as SI, which is suppressed by the RF canceller. Now, as a MIMO transceiver is considered, the SI in the receivers consists of the sum of all the transmit signals, each with their own coupling channel. For the purposes of deriving the signal models in this section, it is assumed that the RF cancellation is performed using the PA output signals, even though that might be unfeasible if the number of transmitters and receivers is large, as is the case in, e.g., massive MIMO devices. However, as shown in **[P6]**, the essential signal model is the same also when using the auxiliary transmitter–based RF canceller illustrated in Fig. 2.5b, and hence this assumption does not affect the generality of the derived signal models. In the linear signal model, the receivers are modeled in the same way as the transmitters, i.e., only the different memory effects are considered, while all the other impairments are represented by a generic error signal. Hence, the signal before the ADC in the ith receiver can be written as:

$$\begin{aligned} y_{i,\mathrm{ADC}}(t) &= f_{i,\mathrm{RX}}(t) \star \left(\sum_{j=1}^{N_\mathrm{t}} h_{ij,\mathrm{SI}}(t) \star x_{j,\mathrm{PA}}(t) - \sum_{j=1}^{N_\mathrm{t}} h_{ij,\mathrm{RFC}}(t) \star x_{j,\mathrm{PA}}(t) \right) + e_{i,\mathrm{RX}}(t) \\ &= f_{i,\mathrm{RX}}(t) \star \sum_{j=1}^{N_\mathrm{t}} \left(h_{ij,\mathrm{SI}}(t) - h_{ij,\mathrm{RFC}}(t) \right) \star x_{j,\mathrm{PA}}(t) + e_{i,\mathrm{RX}}(t) \\ &= \sum_{j=1}^{N_\mathrm{t}} f_{i,\mathrm{RX}}(t) \star h_{ij,\mathrm{RSI}}(t) \star f_{j,\mathrm{TX}}(t) \star x_j(t) \\ &+ \sum_{j=1}^{N_\mathrm{t}} f_{i,\mathrm{RX}}(t) \star h_{ij,\mathrm{RSI}}(t) \star e_{j,\mathrm{TX}}(t) + e_{i,\mathrm{RX}}(t), \end{aligned} \tag{4.2}$$

where $f_{i,\mathrm{RX}}(t)$ is the linear response of the ith receiver, $h_{ij,\mathrm{SI}}(t)$ is the SI coupling channel between the jth transmitter and the ith receiver, $h_{ij,\mathrm{RFC}}(t)$ is the response of

the RF cancellation signal between the jth transmitter and the ith receiver, $h_{ij,\mathrm{RSI}}(t) = h_{ij,\mathrm{SI}}(t) - h_{ij,\mathrm{RFC}}(t)$ is the effective coupling channel after RF cancellation, and $e_{i,\mathrm{RX}}(t)$ is the total noise-plus-modeling-error signal in the ith receiver. Thus, for the purposes of a digital canceller, the RF canceller can in fact be modeled jointly with the propagation channel since it merely adds a certain amount of delayed copies of the PA output signals to the overall received signals, as is evident from (2.2) and (4.2) above. Consequently, the propagation channel and the RF canceller can be modeled by just a single MIMO impulse response $h_{ij,\mathrm{RSI}}(t)$. The same is also true for the auxiliary transmitter–based RF canceller, the only difference being that it utilizes the digital baseband transmit signal instead of the PA output.

Thus, based on (4.2), it can easily be observed that the overall signal model before digital cancellation in the ith receiver can be expressed as follows:

$$y_{i,\mathrm{ADC}}(n) = \sum_{j=1}^{N_\mathrm{t}} \sum_{m=-M_1}^{M_2} h_{ij}^{\mathrm{L}}(m) x_j(n-m) + e_{i,\mathrm{tot}}(n), \tag{4.3}$$

where $h_{ij}^{\mathrm{L}}(n)$ is the total effective linear response between the jth transmitter and the ith receiver, M_1 and M_2 are the numbers of pre-cursor and post-cursor memory taps, respectively, and $e_{i,\mathrm{tot}}(n)$ includes all the unmodeled distortion components, as well as the total noise signal. The pre-cursor taps are introduced here and in the continuation to accurately model the different memory effects occurring in a real IBFD transceiver [**P5, P6**].

4.2.2 Widely Linear Signal Model

Widely linear signal processing, where both the input data and its complex conjugate are utilized when constructing the model, is a useful tool for estimation problems involving noncircular complex data [182], [159, p. 171]. In particular, if the random data cannot be assumed to be circular, the estimation accuracy is typically improved by incorporating also the complex conjugate of the input data into the model. For this reason, as discovered in [**P2**], the tools developed for widely linear systems can also be applied when modeling the residual SI after I/Q imbalance. Namely, then the residual SI signal is noncircular and consequently the system is indeed widely linear [16], as is evident also in (2.5) where both the original transmit signal and its complex conjugate are present in the signal model.

Such a widely linear signal model for IBFD transceivers is presented and described in [**P2**], and in principle it assumes that the SI is only distorted by *(i)* the TX I/Q imbalance, *(ii)* the wireless propagation channel, and *(iii)* the RX I/Q imbalance. The system can thereby be modeled as shown in Fig. 4.3. Stemming from the fact that the possible nonlinearities are not explicitly modeled by the widely linear digital canceller, it is sufficient to again model the amplifiers using linear responses, whose coefficients also include the gain. The noise and the unmodeled distortion components are included in the model via an additive error signal.

Using the model in (2.5) for the TX I/Q mixers, the baseband equivalent of the jth transmit signal is now as follows:

$$x_{j,\mathrm{PA}}(t) = f_{j,\mathrm{TX}}(t) \star \big(k_{j,1,\mathrm{TX}}(t) \star x_j(t) + k_{j,2,\mathrm{TX}}(t) \star x_j^*(t)\big) + e_{j,\mathrm{TX}}(t), \tag{4.4}$$

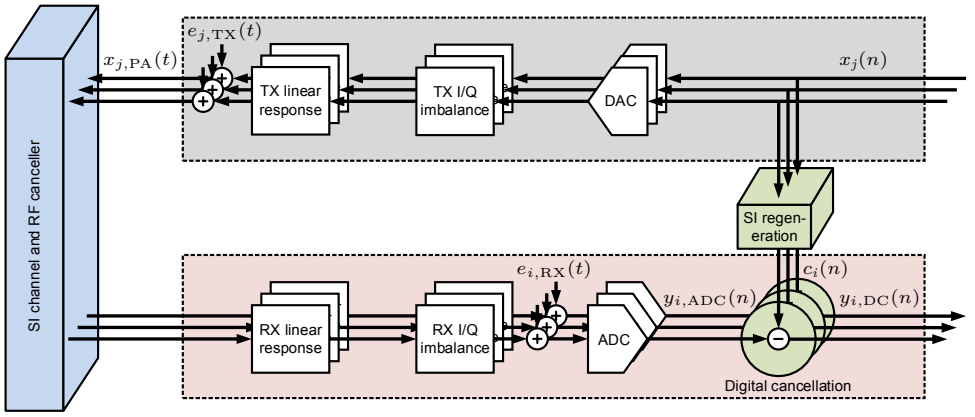


Figure 4.3: The system model used in deriving the widely linear signal model for the digital canceller.

where $k_{j,1,\mathrm{TX}}(t)$ and $k_{j,2,\mathrm{TX}}(t)$ are the responses of the direct and I/Q image components in the jth transmitter, respectively, and the rest of the variables are as defined earlier for the linear system model. The receivers are also assumed to have some I/Q imbalance, modeled again as shown in (2.5), alongside with the linear memory effects. Hence, the signal before the ADC in the ith receiver can be written as:

$$
\begin{aligned}
y_{i,\mathrm{ADC}}(t) &= k_{i,1,\mathrm{RX}}(t) \star f_{i,\mathrm{RX}}(t) \star \sum_{j=1}^{N_\mathrm{t}} \left(h_{ij,\mathrm{SI}}(t) - h_{ij,\mathrm{RFC}}(t)\right) \star x_{j,\mathrm{PA}}(t) \\
&\quad + k_{i,2,\mathrm{RX}}(t) \star f^*_{i,\mathrm{RX}}(t) \star \sum_{j=1}^{N_\mathrm{t}} \left(h^*_{ij,\mathrm{SI}}(t) - h^*_{ij,\mathrm{RFC}}(t)\right) \star x^*_{j,\mathrm{PA}}(t) + e_{i,\mathrm{RX}}(t) \\
&= \sum_{j=1}^{N_\mathrm{t}} \big[k_{i,1,\mathrm{RX}}(t) \star f_{i,\mathrm{RX}}(t) \star h_{ij,\mathrm{RSI}}(t) \star f_{j,\mathrm{TX}}(t) \star k_{j,1,\mathrm{TX}}(t) \\
&\qquad + k_{i,2,\mathrm{RX}}(t) \star f^*_{i,\mathrm{RX}}(t) \star h^*_{ij,\mathrm{RSI}}(t) \star f^*_{j,\mathrm{TX}}(t) \star k^*_{j,2,\mathrm{TX}}(t)\big] \star x_j(t) \\
&\quad + \sum_{j=1}^{N_\mathrm{t}} \big[k_{i,1,\mathrm{RX}}(t) \star f_{i,\mathrm{RX}}(t) \star h_{ij,\mathrm{RSI}}(t) \star f_{j,\mathrm{TX}}(t) \star k_{j,2,\mathrm{TX}}(t) \\
&\qquad + k_{i,2,\mathrm{RX}}(t) \star f^*_{i,\mathrm{RX}}(t) \star h^*_{ij,\mathrm{RSI}}(t) \star f^*_{j,\mathrm{TX}}(t) \star k^*_{j,1,\mathrm{TX}}(t)\big] \star x^*_j(t) \\
&\quad + \sum_{j=1}^{N_\mathrm{t}} \big[k_{i,1,\mathrm{RX}}(t) \star f_{i,\mathrm{RX}}(t) \star h_{ij,\mathrm{RSI}}(t) \star e_{j,\mathrm{TX}}(t) \\
&\qquad + k_{i,2,\mathrm{RX}}(t) \star f^*_{i,\mathrm{RX}}(t) \star h^*_{ij,\mathrm{RSI}}(t) \star e^*_{j,\mathrm{TX}}(t)\big] + e_{i,\mathrm{RX}}(t), \qquad (4.5)
\end{aligned}
$$

where $k_{i,1,\mathrm{RX}}(t)$ and $k_{i,2,\mathrm{RX}}(t)$ are the responses of the direct and I/Q image components in the ith receiver, respectively. This signal is then fed to the ADC, after which, by lumping the various responses together, the final signal model before digital cancellation can be expressed as follows:

$$y_{i,\mathrm{ADC}}(n) = \sum_{j=1}^{N_\mathrm{t}} \sum_{m=-M_1}^{M_2} \left[h_{ij,1}^{\mathrm{WL}}(m) x_j(n-m) + h_{ij,2}^{\mathrm{WL}}(m) x_j^*(n-m) \right] + e_{i,\mathrm{tot}}(n), \quad (4.6)$$

where $h_{ij,1}^{\mathrm{WL}}(n)$ and $h_{ij,2}^{\mathrm{WL}}(n)$ are the total effective responses of the direct and SI image components between the jth transmitter and the ith receiver.

4.2.3 Nonlinear Signal Model

Another approach into modeling the residual SI signal in the digital domain is to assume that the overall distortion is dominated by the nonlinearities produced by the TX PA, as is done, for instance, in **[P4, P5, 18, 125]**, [28]. Then, the essential system model is as shown in Fig. 4.4. In particular, now all the other parts of the transceiver chain are assumed to be ideal, apart from the PAs, meaning that the system is basically a parallel connection of static nonlinearities, followed by linear filters **[P4, P5]**. This type of a signal model is referred to as a PH nonlinearity, as already discussed in some detail in Section 2.4.2.

Now, the jth baseband-equivalent transmit signal is expressed as

$$x_{j,\mathrm{PA}}(t) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} k_{j,p,\mathrm{TX}}(t) \star |x_j(t)|^{p-1} x_j(t) + e_{j,\mathrm{TX}}(t), \quad (4.7)$$

where P is the nonlinearity order, $e_{j,\mathrm{TX}}(t)$ represents again the modeling error and noise of the transmitter, and $k_{j,p,\mathrm{TX}}(t)$ is the response of the pth-order SI term in the jth transmitter, including also the gain. Similar to the earlier cases, the RF canceller can again be modeled jointly with the propagation channel, as also illustrated in Fig. 4.4. In addition, since all the impairments of the RX chains are omitted in this nonlinear signal model, the receivers can be modeled simply as linear filters. Hence, the signal before the analog-to-digital conversion in the ith receiver can simply be expressed as

$$\begin{aligned} y_{i,\mathrm{ADC}}(t) &= f_{i,\mathrm{RX}}(t) \star \sum_{j=1}^{N_\mathrm{t}} \left(h_{ij,\mathrm{SI}}(t) - h_{ij,\mathrm{RFC}}(t) \right) \star x_{j,\mathrm{PA}}(t) + e_{i,\mathrm{RX}}(t) \\ &= \sum_{j=1}^{N_\mathrm{t}} \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} f_{i,\mathrm{RX}}(t) \star h_{ij,\mathrm{RSI}}(t) \star k_{j,p,\mathrm{TX}}(t) \star |x_j(t)|^{p-1} x_j(t) \\ &\quad + \sum_{j=1}^{N_\mathrm{t}} f_{i,\mathrm{RX}}(t) \star h_{ij,\mathrm{RSI}}(t) \star e_{j,\mathrm{TX}}(t) + e_{i,\mathrm{RX}}(t), \end{aligned} \quad (4.8)$$

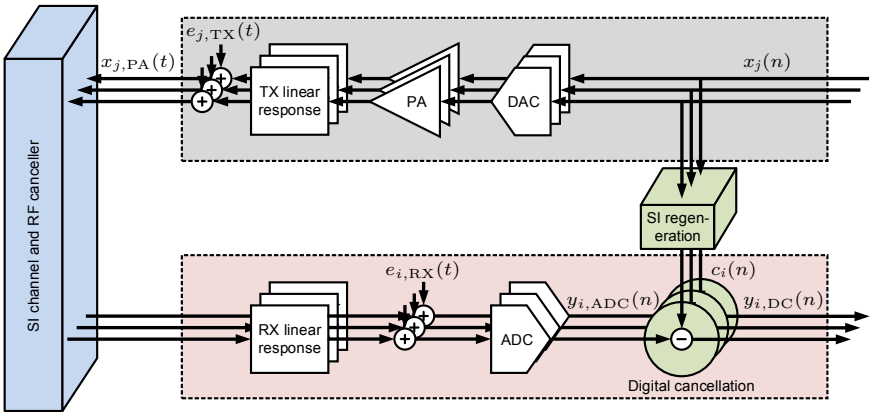


Figure 4.4: The system model used in deriving the nonlinear signal model for the digital canceller.

where the variables are again as defined for the linear and widely linear signal models. The final digital-domain signal can then be written as follows:

$$y_{i,\mathrm{ADC}}(n) = \sum_{j=1}^{N_\mathrm{t}} \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{m=-M_1}^{M_2} h_{ij,p}^{\mathrm{NL}}(m)\psi_p(x_j(n-m)) + e_{i,\mathrm{tot}}(n), \tag{4.9}$$

where $\psi_p(x_j(n)) = |x_j(n)|^{p-1}\, x_j(n)$ is the static pth-order nonlinear basis function, and $h_{ij,p}^{\mathrm{NL}}(m)$ is its total effective response between the jth transmitter and the ith receiver.

4.2.4 Nonlinear Signal Model Incorporating I/Q Imbalance and Transmitter Crosstalk

As was shown already in Chapter 3, under many circumstances both the I/Q imbalance and the PA-induced nonlinear distortion have a significant effect on the overall SI waveform. Hence, there is also a clear need for a signal model that is capable of jointly modeling both of these impairments. Such a signal model is derived in **[P6]** for a MIMO IBFD transceiver, taking also into account the crosstalk between the different transmitters. Especially the crosstalk occurring before the PAs is problematic since it means that the PA input signals are in fact linear combinations of the different transmit signals. This results in a rather complicated transmit signal model, as will be soon demonstrated. However, the crosstalk occurring after the PAs can in fact be considered a part of the wireless coupling channel between the TX and RX antennas since all the TX signals will anyway freely propagate to each receiver. Thus, such post-PA crosstalk is implicitly included also in all the other signal models, even though it is explicitly considered only in this section.

The generic system model for this scenario is illustrated in Fig. 4.5. There, the crosstalk both before and after the nonlinear PAs is shown, alongside with the I/Q imbalance in the transmitters and in the receivers. Adopting again the PH model for

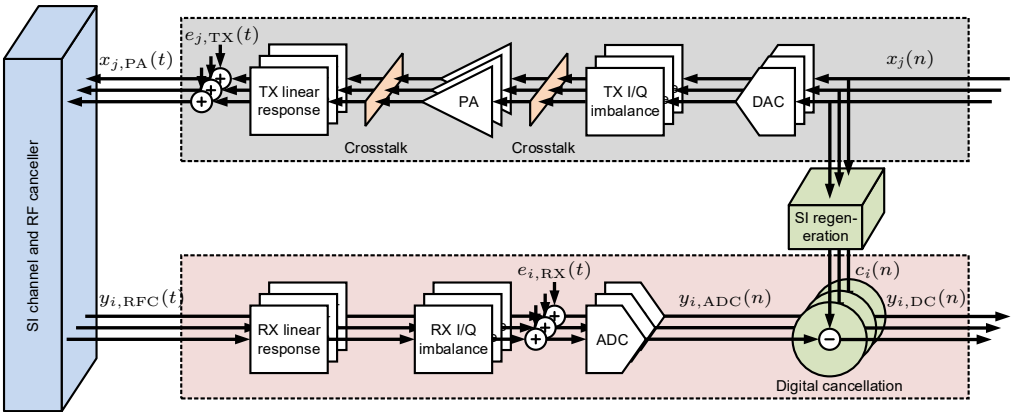


Figure 4.5: The system model used in deriving the nonlinear signal model incorporating crosstalk and I/Q imbalance.

each PA, the baseband-equivalent PA output signal of the jth transmitter can now be written as follows [**P6**]:

$$x_{j,\mathrm{PA}}(t) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{k} k_{j,\boldsymbol{\xi}^k,\mathrm{TX}}(t) \star \prod_{q=1}^{N_\mathrm{t}} x_q(t)^{\xi_q^k} x_q^*(t)^{\xi_{q+N_\mathrm{t}}^k} + e_{j,\mathrm{TX}}(t), \qquad (4.10)$$

where $k_{j,\boldsymbol{\xi}^k,\mathrm{TX}}(t)$ is the response of the different SI terms, N_t is the number of transmit antennas, and $e_{j,\mathrm{TX}}(t)$ is the modeling error plus noise in the jth transmitter. Moreover, $\boldsymbol{\xi}^k$ is the kth combination of the $2N_\mathrm{t} \times 1$ exponent vector $\boldsymbol{\xi}$, and $\|\cdot\|_1$ denotes the L^1-norm. In practice, $\boldsymbol{\xi}$ includes all the different combinations of non-negative integers, whose sum is equal to an odd integer between 1 and P. For more details regarding the above signal model, refer to [**P6**]. Then, the signal after RF cancellation in the ith receiver can be written as follows:

$$\begin{aligned} y_{i,\mathrm{RFC}}(t) &= \sum_{j=1}^{N_\mathrm{t}} \left(h_{ij,\mathrm{SI}}(t) - h_{ij,\mathrm{RFC}}(t)\right) \star x_{j,\mathrm{PA}}(t) \\ &= \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{k} \sum_{j=1}^{N_\mathrm{t}} h_{ij,\mathrm{RSI}}(t) \star k_{j,\boldsymbol{\xi}^k,\mathrm{TX}}(t) \star \prod_{q=1}^{N_\mathrm{t}} x_q(t)^{\xi_q^k} x_q^*(t)^{\xi_{q+N_\mathrm{t}}^k} \\ &\quad + \sum_{j=1}^{N_\mathrm{t}} h_{ij,\mathrm{RSI}}(t) \star e_{j,\mathrm{TX}}(t). \end{aligned} \qquad (4.11)$$

Hence, the essential signal model is still of the same form as in the transmitter output, only with somewhat modified responses, which include the effects of the SI propagation channel and the RF canceller.

Due to the fact that the exponent vector $\boldsymbol{\xi}$ already contains all the combinations under the constraint that the sum of the exponents is equal to the nonlinearity order,

it is clear that the I/Q imbalance in the receivers does not produce any additional SI terms. Namely, since the static SI terms are of the form $\prod_{q=1}^{N_\mathrm{t}} x_q(t)^{a_i} x_q^*(t)^{b_i}$, complex conjugating them will only mean that the corresponding exponent vector changes from $[a_1 \ \cdots \ a_{N_\mathrm{t}} \ \ b_1 \ \cdots \ b_{N_\mathrm{t}}]$ into $[b_1 \ \cdots \ b_{N_\mathrm{t}} \ \ a_1 \ \cdots \ a_{N_\mathrm{t}}]$. However, since the RX input signals already contain all the different combinations of $\boldsymbol{\xi}$, also the latter exponent vector is included, and consequently any RX I/Q imbalance will only result in modified responses or coefficients of the already existing SI terms. Hence, the signal after the ADC in the ith receiver can be expressed as follows:

$$y_{i,\mathrm{ADC}}(n) = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{k \atop \|\boldsymbol{\xi}^k\|_1 = p} \sum_{m=-M_1}^{M_2} h_{i,\boldsymbol{\xi}^k}^{NL}(m) \psi_{\boldsymbol{\xi}^k}(x_1(n-m), \ldots, x_{N_\mathrm{t}}(n-m)) + e_{i,\mathrm{tot}}(n), \tag{4.12}$$

where the static nonlinear basis functions are defined as

$$\psi_{\boldsymbol{\xi}^k}(x_1(n), \ldots, x_{N_\mathrm{t}}(n)) = \prod_{q=1}^{N_\mathrm{t}} x_q(n)^{\xi_q^k} x_q^*(n)^{\xi_{q+N_\mathrm{t}}^k}, \tag{4.13}$$

and $h_{i,\boldsymbol{\xi}^k}^{NL}(m)$ denote their overall effective responses. Note that all the other signal models are in fact special cases of (4.12), and hence it represents the most generic residual SI signal model available in the literature **[P6]**.

4.3 Parameter Estimation and Self-interference Cancellation

In order to actually cancel the SI in the digital domain, the above signal models are used to regenerate the observed residual SI signal. This obviously requires estimating the parameters of the corresponding signal model, for which reason different methods for parameter estimation are presented and discussed in this section. To facilitate more straightforward mathematical derivations, matrix-vector notations are used. Moreover, without loss of generality, the estimation and cancellation procedure is only presented for an individual receiver.

As a starting point, it is assumed that the ith receiver has an observation block of N samples of the residual SI signal $y_{i,\mathrm{ADC}}(n)$ at its disposal for performing the digital cancellation procedure. This signal can be received, for instance, during a calibration period of limited length when there are no other transmissions in the network, as discussed earlier. However, it should be noted that the signal transmitted during this calibration period can also consist of useful data since the transceiver obviously has full knowledge of its own transmit data **[122, 127, 128]**. Without loss of generality, the indexing of the observation block is started from zero for a more illustrative notation, and consequently it is expressed in vector form as follows:

$$\mathbf{y}_{i,\mathrm{ADC}} = \begin{bmatrix} y_{i,\mathrm{ADC}}(0) & y_{i,\mathrm{ADC}}(1) & \cdots & y_{i,\mathrm{ADC}}(N-1) \end{bmatrix}^T. \tag{4.14}$$

The observation block size N is referred to as the *parameter estimation sample size*, and it determines how much data can be used for estimating the SI channel coefficients.

Moreover, the static basis functions of each transmit signal are also collected into a vector, referred to as an *instantaneous basis function vector*. The linear signal model has only one static basis function, i.e., the original transmit signal itself, and hence its instantaneous basis function vector is simply defined as follows:

$$\boldsymbol{\psi}_{\mathrm{L}}(n) = \begin{bmatrix} x_1(n) & x_2(n) & \cdots & x_{N_{\mathrm{t}}}(n) \end{bmatrix}. \tag{4.15}$$

Denoting the number of static basis functions for all transmit signals by K, in this case $K = N_{\mathrm{t}}$. Considering then the widely linear signal model, the corresponding instantaneous basis function vector is defined as

$$\boldsymbol{\psi}_{\mathrm{WL}}(n) = \begin{bmatrix} \boldsymbol{\psi}_{\mathrm{L}}(n) & \boldsymbol{\psi}_{\mathrm{L}}^{*}(n) \end{bmatrix}, \tag{4.16}$$

and now $K = 2N_{\mathrm{t}}$. That is, the widely linear signal model has two static basis functions for each transmit signal: the original signal and its complex conjugate. The nonlinear signal model, on the other hand, has $K = N_{\mathrm{t}}\frac{P+1}{2}$ static basis functions in total and the corresponding instantaneous basis function vector is expressed as

$$\boldsymbol{\psi}_{\mathrm{NL}}(n) = \begin{bmatrix} \psi_1(x_1(n)) & \psi_3(x_1(n)) & \cdots & \psi_P(x_1(n)) & \psi_1(x_2(n)) & \cdots & \psi_P(x_{N_{\mathrm{t}}}(n)) \end{bmatrix}, \tag{4.17}$$

where $\psi_p(x_j(n))$ is the static pth-order nonlinear basis function, defined in Section 4.2.3. Finally, the instantaneous basis function vector of the nonlinear signal model with crosstalk and I/Q imbalance is defined as follows:

$$\boldsymbol{\psi}_{\mathrm{NLC}}(n) = \begin{bmatrix} \psi_{\boldsymbol{\xi}^1}(\boldsymbol{\psi}_{\mathrm{L}}(n)) & \psi_{\boldsymbol{\xi}^2}(\boldsymbol{\psi}_{\mathrm{L}}(n)) & \cdots & \psi_{\boldsymbol{\xi}^K}(\boldsymbol{\psi}_{\mathrm{L}}(n)) \end{bmatrix}, \tag{4.18}$$

where $\psi_{\boldsymbol{\xi}^k}(\boldsymbol{\psi}_{\mathrm{L}}(n))$ is as defined in (4.13) but just with the different arguments lumped into a single vector. It is shown in [**P6**] that the value of K is in this case given by

$$K = \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \binom{p + 2N_{\mathrm{t}} - 1}{2N_{\mathrm{t}} - 1}. \tag{4.19}$$

Note that this signal model has no basis functions corresponding to an individual transmit signal since each $\psi_{\boldsymbol{\xi}^k}(\boldsymbol{\psi}_{\mathrm{L}}(n))$ already contains a contribution from all the transmit signals, as defined earlier. This stems from the crosstalk, which results in each transmitter essentially transmitting a linear combination of the N_{t} transmit signals. Hence, all the static basis functions for all the transmit signals can be generated by just choosing the right configuration of the exponent vector $\boldsymbol{\xi}$ from the set of K possibilities.

The total number of basis functions for all the signal models is then simply KM, where $M = M_1 + M_2 + 1$ is the total amount of memory. That is, the term *basis function* is used to refer to all the KM entities that include also the delayed versions of the static or instantaneous basis functions. This terminology is adopted throughout the thesis.

It should also be noted here that, when applying the nonlinear basis functions to generate the nonlinear SI terms for cancellation purposes, a higher sampling frequency should be used to avoid aliasing [**124**]. This stems from the fact that the bandwidth of the pth-order nonlinearity is p times the bandwidth of the original signal [173]. Consequently, if the input signal of a nonlinear basis function is not properly oversampled, some of the nonlinear distortion will alias onto the original signal band, resulting in an inaccurate model of the true nonlinear process. Hence, in theory, the input signal of a 3rd-order basis function must be oversampled by a factor of 3 to avoid aliasing at the output, and so forth. In practice, however, the higher order nonlinearities have typically very little spectral content close to the edge of their theoretical bandwidth, and thus less oversampling usually suffices [**124**]. What is more, since in this case only the inband content of the generated nonlinear terms actually matters, some aliasing outside the actual signal band can be tolerated.

After generating the basis functions with sufficient oversampling, the resulting nonlinear signal is then decimated back to the original sampling frequency, with appropriate filtering to avoid aliasing effects. For notational simplicity, the necessary interpolation and decimation procedures are assumed to be implicitly included in the respective basis functions in the derivations of this chapter, and hence they are not explicitly considered. Equivalently, it can be assumed that the signals have already been sufficiently oversampled for the considered nonlinearity orders. For further discussion regarding this aspect, refer to [**124**].

Having now defined the necessary signal vectors, different methods for calculating the SI channel estimate are presented in the following sections. The presented estimation and cancellation algorithms are then evaluated in Chapter 5 with both simulations and measurements.

4.3.1 Block Least Squares–Based Estimation and Cancellation

The least squares (LS) algorithm is a powerful and versatile parameter estimation tool since it makes no implicit or explicit assumptions regarding the statistics of the total noise signal [109, p. 219]. This is beneficial in the context of digital SI cancellation since the noise cannot in general be expected to follow any particular probability distribution due to the various RF impairments. For this reason, this section presents an LS-based digital cancellation algorithm, which operates in a block-wise nature. That is, it estimates the SI channel coefficients for a block of N received samples, after which these estimated coefficients can be used to cancel the SI until the SI channel estimation procedure is again repeated after a certain period of time. The basic operating principle of the LS-based canceller is shown on a general level in Fig. 4.6.

To lay out the LS estimation procedure in detail, let us first express the residual SI signal in the ith receiver using the defined vectors as follows:

$$y_{i,\mathrm{ADC}}(n) = \sum_{m=-M_1}^{M_2} \boldsymbol{\psi}(n-m)\mathbf{h}_i(m) + e_{i,\mathrm{tot}}(n), \tag{4.20}$$

where $n = 0, 1, \ldots, N$, $\mathbf{h}_i(m)$ is the $(M_1 + M_2 + 1) \times 1$ channel coefficient vector corresponding to the basis functions with lag m, and $\boldsymbol{\psi}(n)$ is an instantaneous basis function

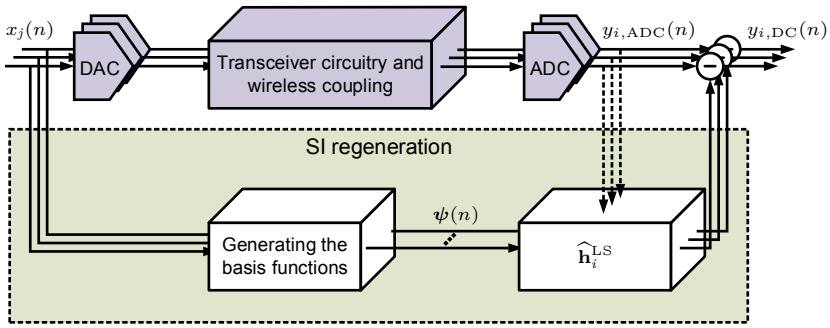


Figure 4.6: A generic illustration of the LS-based digital SI canceller.

vector without a subscript associating it with any particular signal model to present the parameter estimation procedure in a generic fashion. That is, it is either as shown in (4.15), (4.16), (4.17), or (4.18), depending on the utilized signal model. The vectors $\boldsymbol{\psi}(n-m)$ constitute then the complete set of basis functions when considering all the values of m. Note that (4.20) implicitly assumes that the receiver can use transmit data also from outside the given block of N samples for cancellation processing. This assumption is well justified from an estimation theoretic perspective since the own transmit signal is obviously known within the device and hence it does not bring any new information into the system, unlike the observed residual SI signal.

To express the whole observation block of the residual SI signal, the instantaneous basis function vectors can be collected into a single *convolution data matrix* [**P2**], which is defined as follows:

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\psi}(M_1) & \boldsymbol{\psi}(M_1-1) & \cdots & \boldsymbol{\psi}(-M_2) \\ \boldsymbol{\psi}(M_1+1) & \boldsymbol{\psi}(M_1) & \cdots & \boldsymbol{\psi}(-M_2+1) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\psi}(N+M_1-1) & \boldsymbol{\psi}(N+M_1-2) & \cdots & \boldsymbol{\psi}(N-M_2-1) \end{bmatrix}. \tag{4.21}$$

The corresponding complete channel vector for the ith receiver is then simply

$$\mathbf{h}_i = \begin{bmatrix} \mathbf{h}_i^T(-M_1) & \mathbf{h}_i^T(-M_1+1) & \cdots & \mathbf{h}_i^T(M_2) \end{bmatrix}^T, \tag{4.22}$$

which contains the coefficients of all the basis functions. With these, the residual SI vector in the ith receiver can be written as follows:

$$\mathbf{y}_{i,\mathrm{ADC}} = \boldsymbol{\Psi}\mathbf{h}_i + \mathbf{e}_i, \tag{4.23}$$

where $\mathbf{e}_i$ is the $N \times 1$ noise-plus-modeling-error vector.

Since the overall signal model in (4.23) is in fact linear in parameters with both the received signal vector $\mathbf{y}_{i,\mathrm{ADC}}$ and the matrix $\boldsymbol{\Psi}$ being obviously known, linear LS can be used to estimate the channel coefficients. Assuming that $\boldsymbol{\Psi}$ has full column rank, the LS solution to the SI channel estimation problem is given by [**P2, 18**], [109, p. 223]:

$$\widehat{\mathbf{h}}_i^{\mathrm{LS}} = \min_{\mathbf{h}_i} \|\mathbf{y}_{i,\mathrm{ADC}} - \boldsymbol{\Psi}\mathbf{h}_i\|_2^2 = \left(\boldsymbol{\Psi}^H\boldsymbol{\Psi}\right)^{-1}\boldsymbol{\Psi}^H\mathbf{y}_{i,\mathrm{ADC}}, \tag{4.24}$$

where $\|\cdot\|_2$ denotes the L^2-norm. The accuracy of this estimate is largely determined by the parameter estimation sample size N, as is well known in estimation theory [109]. However, the computational complexity of the estimation procedure is higher for larger values of N, and hence a proper trade-off between accuracy and complexity must be determined.

The digital cancellation signal is then obtained by applying the estimated channel coefficients in $\widehat{\mathbf{h}}_i^{\mathrm{LS}}$ into the corresponding SI terms. This is done by first stacking the instantaneous basis function vectors to form the complete basis function vector as follows:

$$\boldsymbol{\Psi}(n) = \begin{bmatrix} \boldsymbol{\psi}(n+M_1) & \boldsymbol{\psi}(n+M_1-1) & \cdots & \boldsymbol{\psi}(n-M_2) \end{bmatrix}, \tag{4.25}$$

which is of the same form as the nth row of the convolution data matrix $\boldsymbol{\Psi}$. The corresponding digital cancellation signal in the ith receiver is then simply

$$c_i(n) = \boldsymbol{\Psi}(n)\widehat{\mathbf{h}}_i^{\mathrm{LS}}. \tag{4.26}$$

Thus, the signal after the LS-based digital canceller is as follows:

$$y_{i,\mathrm{DC}}(n) = y_{i,\mathrm{ADC}}(n) - c_i(n) = y_{i,\mathrm{ADC}}(n) - \boldsymbol{\Psi}(n)\widehat{\mathbf{h}}_i^{\mathrm{LS}} \approx e_{i,\mathrm{tot}}(n), \tag{4.27}$$

where the final approximation is obviously valid only when the SI channel estimate is sufficiently accurate.

An important aspect of the LS estimation scheme is its block-wise nature. Namely, the SI channel estimate is always calculated for a certain piece of the observed signal, whose length in this case is determined by the parameter estimation sample size N. Thus, LS is well suited for scenarios where the SI channel is estimated during a dedicated training period when there are no other transmissions in the network [**122, 128**]. The same estimate must then be used until the next calibration period, during which the LS estimate is recalculated. Moreover, it should be noted that obtaining the LS estimate requires solving the pseudoinverse of the convolution data matrix $\boldsymbol{\Psi}$, as can be observed in (4.24). This can be computationally intensive, especially if the number of basis functions is large. Hence, especially in mobile-scale devices, simpler and less computationally demanding estimation methods might be preferable.

4.3.2 Least Mean Squares–Based Adaptive Estimation and Cancellation

One approach to decreasing the computational requirements of the digital SI cancellation procedure is to utilize adaptive algorithms for estimating the channel coefficients. A widely used solution for this type of an estimation problem is the LMS parameter learning algorithm, which has been described and evaluated in [**P4, P5, 125**]. The basic operating principle of the LMS-based digital canceller is illustrated in Fig. 4.7. Essentially, the LMS canceller aims at minimizing the power of its output signal [85, p. 150], i.e., $y_{i,\mathrm{DC}}(n)$, which it does by utilizing a predetermined signal model for the residual SI.

However, as also shown in Fig. 4.7, before the actual parameter learning, the different static basis functions must be orthogonalized. The reason for this is the poor convergence

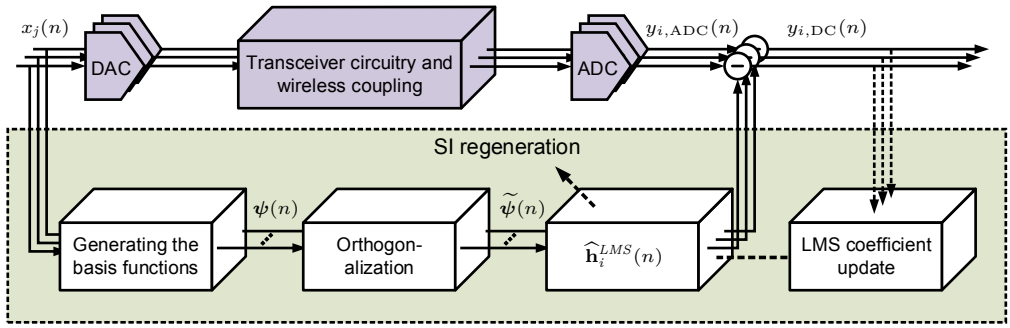


Figure 4.7: A generic illustration of the LMS-based digital SI canceller.

performance of the LMS algorithm if the elements of the input vector are highly correlated, caused by the large eigenvalue spread of the input signal covariance matrix [**125**], [84, p. 417], [66]. Since in this case the static basis functions can indeed be expected to be correlated as they are all dependent on the original transmit signal, orthogonalizing them is necessary to ensure efficient parameter learning by the LMS algorithm. Although there are various alternative methods for performing the orthogonalization, in this thesis it is done with an orthogonalization matrix that can be obtained starting from the covariance matrix of the instantaneous basis functions, defined as follows [84, p. 100]:

$$\mathbf{R}_{\boldsymbol{\psi}} = \mathbb{E}\left[\boldsymbol{\psi}^H(n)\boldsymbol{\psi}(n)\right], \tag{4.28}$$

where the subscript of the instantaneous basis function vector has again been omitted to present the LMS-based solution in a generic fashion. The next step in deriving the orthogonalization matrix is calculating the eigendecomposition of the above covariance matrix:

$$\mathbf{R}_{\boldsymbol{\psi}} = \mathbf{U}_{\boldsymbol{\psi}}\boldsymbol{\Lambda}_{\boldsymbol{\Psi}}\mathbf{U}_{\boldsymbol{\psi}}^H, \tag{4.29}$$

where $\boldsymbol{\Lambda}_{\boldsymbol{\Psi}}$ is a diagonal matrix containing the eigenvalues of $\mathbf{R}_{\boldsymbol{\psi}}$, and $\mathbf{U}_{\boldsymbol{\psi}}$ is a unitary matrix containing the corresponding eigenvectors [100, p. 21]. The orthogonalization matrix is then simply given by [**125**], [100, p. 140]

$$\mathbf{S}_{\boldsymbol{\psi}} = \mathbf{U}_{\boldsymbol{\psi}}\boldsymbol{\Lambda}_{\boldsymbol{\Psi}}^{-1/2}, \tag{4.30}$$

where $\boldsymbol{\Lambda}_{\boldsymbol{\Psi}}^{-1/2}$ denotes an element-wise square root and inverse of the diagonal elements. The orthogonalization matrix is written here in a slightly different form than in [**125**] since now the instantaneous basis functions are expressed as a row vector, whereas in [**125**] they are collected into a column vector. Nevertheless, the orthogonalization principle is still identical, apart from these notational differences. The static basis functions can then be orthogonalized simply by

$$\widetilde{\boldsymbol{\psi}}(n) = \boldsymbol{\psi}(n)\mathbf{S}_{\boldsymbol{\psi}}, \tag{4.31}$$

which can easily be shown to be orthogonal as follows:

$$\mathbb{E}\left[\widetilde{\boldsymbol{\psi}}^H(n)\widetilde{\boldsymbol{\psi}}(n)\right] = \mathbf{S}_{\boldsymbol{\psi}}^H\mathbb{E}\left[\boldsymbol{\psi}^H(n)\boldsymbol{\psi}(n)\right]\mathbf{S}_{\boldsymbol{\psi}} = \boldsymbol{\Lambda}_{\boldsymbol{\Psi}}^{-1/2}\mathbf{U}_{\boldsymbol{\psi}}^H\mathbf{U}_{\boldsymbol{\psi}}\boldsymbol{\Lambda}_{\boldsymbol{\Psi}}\mathbf{U}_{\boldsymbol{\psi}}^H\mathbf{U}_{\boldsymbol{\psi}}\boldsymbol{\Lambda}_{\boldsymbol{\Psi}}^{-1/2} = \mathbf{I}_K, \tag{4.32}$$

where $\mathbf{I}_K$ is a $K \times K$ identity matrix. Note that the orthogonalization matrix $\mathbf{S}_{\boldsymbol{\psi}}$ only depends on the statistical properties of the original transmit signal (via the covariance matrix $\mathbf{R}_{\boldsymbol{\psi}}$) and hence it does not change with respect to time, as long as the transmit waveforms remains the same. This means that the orthogonalization matrix can be precomputed offline, and only the actual orthogonalization in (4.31) must be performed in real time. However, it is also possible to calculate the orthogonalization matrix adaptively during the actual digital cancellation procedure, as shown in **[P5]**.

Having orthogonalized the basis functions, they can then be used for learning the SI channel coefficients with the LMS algorithm. Now, the input vector of the LMS filter, containing all the orthogonalized basis functions, is defined as follows:

$$\widetilde{\boldsymbol{\Psi}}(n) = \begin{bmatrix} \widetilde{\boldsymbol{\psi}}(n+M_1) & \widetilde{\boldsymbol{\psi}}(n+M_1-1) & \cdots & \widetilde{\boldsymbol{\psi}}(n-M_2) \end{bmatrix}. \tag{4.33}$$

Then, denoting the LMS SI channel estimate in the ith receiver after n iterations by $\widehat{\mathbf{h}}_i^{\mathrm{LMS}}(n)$ (where $n = 0, 1, \ldots, N$), the cancelled signal is given by:

$$y_{i,\mathrm{DC}}(n) = y_{i,\mathrm{ADC}}(n) - \widetilde{\boldsymbol{\Psi}}(n)\widehat{\mathbf{h}}_i^{\mathrm{LMS}}(n), \tag{4.34}$$

after which the LMS algorithm updates the SI channel estimate using the following rule **[125]**:

$$\widehat{\mathbf{h}}_i^{\mathrm{LMS}}(n+1) = \widehat{\mathbf{h}}_i^{\mathrm{LMS}}(n) + \mathbf{M} y_{i,\mathrm{DC}}(n) \widetilde{\boldsymbol{\Psi}}^H(n), \tag{4.35}$$

where $\mathbf{M}$ is a diagonal matrix containing the step sizes for the different orthogonalized basis functions on its diagonal. If no further side information is available, the channel estimate is initialized as $\widehat{\mathbf{h}}_i^{\mathrm{LMS}}(0) = \mathbf{0}$.

Comparing the LS-based and LMS-based channel estimation procedures, it can be observed that the LMS rule in (4.35) requires only additions and multiplications, whereas the LS estimation involves a costly matrix inversion, among other matrix operations. Hence, as will be shown in Section 4.3.4, the LMS algorithm is computationally more efficient than performing the estimation with LS. Moreover, the LMS-based digital canceller is also capable of tracking the SI channel under time-varying conditions, unlike the LS estimator which assumes the SI channel to be static during the whole observation period of N samples. However, as is well known in estimation theory, with a sufficient amount of learning data and Gaussian-distributed noise, the variance of the channel estimate given by the LMS algorithm is higher than that given by the LS estimator [84, p. 397]. This is an inherent cost for the many upsides of the LMS canceller. Consequently, in some cases a suitable compromise might be to utilize a parameter estimation solution that falls somewhere in between LS and LMS in terms of accuracy and complexity, such as the recursive least squares (RLS) algorithm [84, p. 562].

4.3.3 Model Complexity Reduction with Principal Component Analysis

Under some circumstances, it might be desirable to reduce the signal model complexity by omitting the basis functions that are insignificant in terms of the modeling accuracy. Namely, while reducing the amount of basis functions can be expected to decrease the

overall computational complexity of most estimation and cancellation procedures, it also results in a lower variance for the obtained SI channel estimate due to the smaller number of parameters [109], which improves the cancellation performance. These aspects are especially crucial for the signal model incorporating crosstalk and I/Q imbalance as it contains a large amount of basis functions, many of which are negligibly weak [**P6**]. What is more, there are very few heuristic techniques for determining the insignificant basis functions in this signal model, unlike in the simple nonlinear signal model where the higher-order basis functions tend to be also less significant than the ones of lower order. Hence, this calls for an automated approach to identifying the insignificant basis functions.

In [**P6**] a PCA-based approach is proposed, which only retains the desired number of basis functions in the signal model. Essentially, the PCA[1] forms linear combinations of the original basis functions that account for as much of the total variance in the observed SI as possible [85]. In practice, this transformation is done with a matrix multiplication where the transformation matrix is the outcome of the PCA. Furthermore, these new transformed linear combinations, or components, are ordered in terms of their variances such that the first one accounts for most of the total variance of the observed signal, while the last one is the least significant. Hence, by neglecting a given number of components starting from the last one, the corresponding amount of basis functions is removed with the smallest possible effect on the modeling accuracy. In this sense, the PCA-based approach allows for a simple controlling mechanism of the signal model complexity, in much the same way as the nonlinearity order can be used to adjust the accuracy and complexity of the simple nonlinear signal model. There are also works available where similar complexity reduction schemes have been implemented with other techniques, such as compressive sensing (sometimes referred to as compressive sampling) [1, 81, 164]. Nevertheless, in this thesis, PCA is adopted since it is a straightforward method for the complexity reduction of the proposed signal model while also providing almost the same performance as compressive sensing when high modeling accuracy is required [1].

As a starting point for the PCA-based model complexity reduction, the SI channel coefficients for each receiver are first estimated with the complete and unreduced signal model. This estimate can be calculated using either LS or LMS as described above. During this initial channel estimation stage, the highest possible transmit power should be used as then the truly insignificant basis functions can be reliably detected. Namely, the nonlinear terms that are insignificantly weak with the highest possible transmit power will be even weaker with any lower transmit power. In the case where $N_\mathrm{r} > 1$, the different channel estimates are averaged since only the relative magnitudes of the different basis functions are of significance. This also means that there is no need to have a separate procedure for controlling the model complexity in the different receivers. Denoting this initial channel estimate by $\widehat{\mathbf{h}}_0$, the different basis functions are then weighted by their corresponding channel coefficients to establish their true magnitude in the total signal model. The hereby resulting weighted data matrix is as follows:

$$\mathbf{\Psi}_0 = \left(\mathbf{1}_N \widehat{\mathbf{h}}_0^T\right) \circ \mathbf{\Psi}, \tag{4.36}$$

[1] In some contexts, PCA is also referred to as the discrete Karhunen-Loève transform [85, p. 418].

where $\mathbf{1}_N$ is a column vector consisting of N ones. Note that the parameter estimation sample size used in calculating the PCA might be different than that used for estimating the SI channel coefficients, but for notational simplicity the same variable (N) is used here for both. Hence, the convolution data matrix $\mathbf{\Psi}$ used to obtain the weighted data matrix $\mathbf{\Psi}_0$ for PCA processing is as defined in (4.21). However, if using the LMS algorithm for calculating the initial channel estimate $\widehat{\mathbf{h}}_0$, then $\mathbf{\Psi}$ should be constructed using the orthogonalized basis functions, since these basis functions are also used when obtaining the SI channel estimate.

The actual PCA transformation matrix is then determined by calculating the singular value decomposition (SVD) of the matrix $\mathbf{\Psi}_0$ as follows:

$$\mathbf{\Psi}_0 = \mathbf{U}_0 \mathbf{\Sigma}_0 \mathbf{V}_0^H, \tag{4.37}$$

where $\mathbf{U}_0$ and $\mathbf{V}_0$ are matrices containing the left and right singular vectors, respectively, and $\mathbf{\Sigma}_0$ is a diagonal matrix consisting of the singular values, ordered based on their magnitudes. The PCA transformation matrix is then simply

$$\mathbf{S}_{\text{PCA}} = \mathbf{V}_0 \mathbf{\Sigma}_0^{-1}, \tag{4.38}$$

where the diagonal values of $\mathbf{\Sigma}_0$ are inverted to normalize the magnitudes of the resulting transformed basis functions [**P6**], although this complexity reduction scheme will also function without the normalization. Then, based on the earlier assumption regarding the ordering of the singular values, the columns of the transformation matrix $\mathbf{S}_{\text{PCA}}$ correspond to the new transformed basis functions in decreasing order of significance [**P6**]. This means that, by removing a given number of the rightmost columns of $\mathbf{S}_{\text{PCA}}$, the number of basis functions in the signal model can be reduced with minimal effect on the modeling accuracy. Denoting the number of included basis functions by L, the reduced and final transformation matrix is thereby defined as follows:

$$\widetilde{\mathbf{S}}_{\text{PCA}} = \begin{bmatrix} \mathbf{s}_{\text{PCA},1} & \mathbf{s}_{\text{PCA},2} & \cdots & \mathbf{s}_{\text{PCA},L} \end{bmatrix}, \tag{4.39}$$

where $\mathbf{s}_{\text{PCA},n}$ is the nth column of $\mathbf{S}_{\text{PCA}}$. Finally, the reduced basis function vector is simply calculated as

$$\breve{\mathbf{\Psi}}(n) = \mathbf{\Psi}(n) \widetilde{\mathbf{S}}_{\text{PCA}}, \tag{4.40}$$

or as $\breve{\mathbf{\Psi}}(n) = \widetilde{\mathbf{\Psi}}(n) \widetilde{\mathbf{S}}_{\text{PCA}}$ if LMS has been used to obtain the initial channel estimate. This reduced vector can then be used to either form the convolution data matrix for LS estimation, or it can be used directly for LMS parameter learning by replacing $\widetilde{\mathbf{\Psi}}(n)$ in (4.34) and (4.35).

It should be emphasized that the PCA transformation matrix is calculated only once, after which it can be repeatedly used to reduce the number of basis functions. This means that the only procedure that must be performed during the actual SI cancellation is the matrix multiplication in (4.40), while the transformation matrix itself can be computed offline. However, considering that the PCA is performed for the whole data matrix, including the channel taps, the transformation matrix might have to be recomputed if the surrounding environment around the IBFD transceiver changes drastically. Namely,

in that case, some of the insignificant memory taps might become stronger and hence omitting them would result in a non-negligible modeling error. The numerical evaluation of this PCA-based complexity reduction scheme is performed in Section 5.2, where it is applied to the nonlinear signal model incorporating crosstalk and I/Q imbalance.

4.3.4 Computational Complexity of Digital Cancellation

Let us then briefly analyze the computational complexities of the two alternative parameter estimation algorithms for an individual receiver. Here, the so-called *Big O notation*, denoted by $\mathcal{O}(\cdot)$, is used to characterize their asymptotic complexities for large data sets, written with respect to complex arithmetic operations [118, p. 107]. Such analysis describes how the number of arithmetic operations required for large data sets is related to the dimensions of the input matrices and/or vectors, which is a common approach for comparing the computational complexities of different algorithms [55, 116, 248].

Least Squares

Starting from the unreduced LS-based solution, it consists of first estimating the SI channel coefficients with (4.24), after which the SI signal is regenerated and cancelled. The former consists of the following parts:

- calculating the matrix product $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$;
- calculating the matrix-vector product $\boldsymbol{\Psi}^H\mathbf{y}_{i,\mathrm{ADC}}$;
- inverting the matrix $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$;
- calculating the matrix-vector product between the inverse of $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ and $\boldsymbol{\Psi}^H\mathbf{y}_{i,\mathrm{ADC}}$.

The consecutive SI regeneration and cancellation simply consists of first calculating the matrix-vector product $\boldsymbol{\Psi}(n)\widehat{\mathbf{h}}_i^{\mathrm{LS}}$ and then subtracting it from the corresponding input sample $y_{i,\mathrm{ADC}}(n)$. In order to quantify then the overall arithmetic complexity, recall that $\mathbf{y}_{i,\mathrm{ADC}}$ is a $N\times 1$ vector, $\widehat{\mathbf{h}}_i^{\mathrm{LS}}$ is a $KM\times 1$ vector, and $\boldsymbol{\Psi}$ is a $N\times KM$ matrix, where N is the parameter estimation sample size, K is the number of static basis functions, and M is the total number of memory taps. To facilitate a straightforward comparison with the LMS-based solution, the cancellation is assumed to be performed over N samples, meaning that the regeneration and cancellation must be repeated N times.

It is easy to show that the arithmetic complexity of calculating the matrix product $\boldsymbol{\Psi}^H\boldsymbol{\Psi}$ between the $KM\times N$ matrix and the $N\times KM$ matrix is $\mathcal{O}\big(K^2M^2N\big)$, while the corresponding complexity of inverting the resulting $KM\times KM$ matrix is $\mathcal{O}\big(K^3M^3\big)$. The latter assumes that a LUP decomposition–based approach is used to find the inverse [52, p. 828]. Determining then the arithmetic complexities of the rest of the operations in a similar manner, the total complexity of the LS-based digital cancellation procedure is

$$\begin{aligned}&\mathcal{O}\big(K^2M^2N\big)+\mathcal{O}(KMN)+\mathcal{O}\big(K^3M^3\big)+\mathcal{O}\big(K^2M^2\big)+\mathcal{O}(KMN)\\&=\mathcal{O}\big(K^3M^3\big)+\mathcal{O}\big(K^2M^2N\big)\,.\end{aligned}\tag{4.41}$$

Furthermore, in any practical system $N\gg KM$, which means that the term $\mathcal{O}\big(K^2M^2N\big)$ typically dominates the arithmetic complexity asymptotically.

Considering then the arithmetic complexity when utilizing the PCA-based basis function reduction scheme, the estimation and cancellation procedure itself remains the same, apart from two additional matrix/vector multiplications. In particular, now the data matrix $\mathbf{\Psi}$ is first multiplied with the $KM \times L$ reduced transformation matrix $\widetilde{\mathbf{S}}_{\mathrm{PCA}}$, which decreases the number of coefficients from KM to L. Moreover, also the basis function vector $\mathbf{\Psi}(n)$ must be multiplied with the same matrix before SI regeneration and cancellation. Following then an identical derivation as done for the unreduced signal model above, it can easily be shown that now the arithmetic complexity is dominated by the term $\mathcal{O}(KMLN)$. Hence, as obviously $L < KM$, the PCA-based model complexity reduction scheme reduces the asymptotic computational complexity of the LS-based digital canceller.

Least Mean Squares

Analyzing next the adaptive LMS-based digital cancellation procedure, in the unreduced case each iteration involves the following computations:

- calculating the vector-matrix product $\boldsymbol{\psi}(n)\mathbf{S}_{\boldsymbol{\psi}}$;
- calculating the dot product $\widetilde{\mathbf{\Psi}}(n)\widehat{\mathbf{h}}_i^{\mathrm{LMS}}(n)$, and subtracting it from $y_{i,\mathrm{ADC}}(n)$;
- calculating the matrix-vector product $\mathbf{M}y_{i,\mathrm{DC}}(n)\widetilde{\mathbf{\Psi}}^H(n)$ and adding it to $\widehat{\mathbf{h}}_i^{\mathrm{LMS}}(n)$.

Since the complete signal model has K static basis functions and M memory taps, it is easy to show that the overall asymptotic arithmetic complexity over N iterations is now

$$\mathcal{O}\big(K^2N\big) + \mathcal{O}(KMN) + \mathcal{O}(KMN) = \mathcal{O}\big(K^2N\big) + \mathcal{O}(KMN)\,. \tag{4.42}$$

Noting then that typically the number of memory taps is higher than the number of static basis functions, i.e., $M > K$, it can be concluded that the term $\mathcal{O}(KMN)$ usually dominates asymptotically.

When using the PCA-based complexity reduction scheme in conjunction with the LMS-based parameter learning and cancellation, the only alteration to the procedure with the unreduced signal model is replacing the orthogonalization with the PCA transformation. In particular, instead of calculating the vector-matrix product $\boldsymbol{\psi}(n)\mathbf{S}_{\boldsymbol{\psi}}$, the complete $1 \times KM$ basis function vector is multiplied with the reduced $KM \times L$ transformation matrix $\widetilde{\mathbf{S}}_{\mathrm{PCA}}$, which decreases the number of coefficients from KM to L. It can then easily be shown that in this case the arithmetic complexity is dominated by the term $\mathcal{O}(KMLN)$, meaning that the PCA-based model complexity reduction scheme somewhat increases the arithmetic complexity of the LMS-based digital canceller. Consequently, in this case, the primary benefit of reducing the number of basis functions is the increased estimation accuracy and faster convergence, thanks to the smaller number of coefficients that must be estimated.

To facilitate a more straightforward comparison between the computational requirements of the LS-based and LMS-based digital cancellers, as well as to determine the relative complexities of the different signal models, Table 4.1 shows the arithmetic complexities for all combinations of signal models and parameter estimation algorithms. The expressions listed in the table have been obtained simply by substituting K with the number of static basis functions in the corresponding signal model and determining

Table 4.1: The asymptotic arithmetic complexity of digital cancellation with different signal models and parameter estimation algorithms.

	Linear signal model	**Widely linear signal model**	**Nonlinear signal model**	**Nonlinear signal model incorporating crosstalk and I/Q imbalance**
LS without PCA-based complexity reduction	$\mathcal{O}\big(N_\mathrm{t}^2M^2N\big)$	$\mathcal{O}\big(N_\mathrm{t}^2M^2N\big)$	$\mathcal{O}\big(N_\mathrm{t}^2P^2M^2N\big)$	$\mathcal{O}\Big(\binom{P+2N_\mathrm{t}-1}{2N_\mathrm{t}-1}^2M^2N\Big)$
LS with PCA-based complexity reduction	$\mathcal{O}(N_\mathrm{t}MLN)$	$\mathcal{O}(N_\mathrm{t}MLN)$	$\mathcal{O}(N_\mathrm{t}PMLN)$	$\mathcal{O}\Big(\binom{P+2N_\mathrm{t}-1}{2N_\mathrm{t}-1}MLN\Big)$
LMS without PCA-based complexity reduction	$\mathcal{O}(N_\mathrm{t}MN)$	$\mathcal{O}(N_\mathrm{t}MN)$	$\mathcal{O}\big(N_\mathrm{t}^2P^2N\big)$ $+\mathcal{O}(N_\mathrm{t}PMN)$	$\mathcal{O}\Big(\binom{P+2N_\mathrm{t}-1}{2N_\mathrm{t}-1}^2N\Big)$ $+\mathcal{O}\Big(\binom{P+2N_\mathrm{t}-1}{2N_\mathrm{t}-1}MN\Big)$
LMS with PCA-based complexity reduction	$\mathcal{O}(N_\mathrm{t}MLN)$	$\mathcal{O}(N_\mathrm{t}MLN)$	$\mathcal{O}(N_\mathrm{t}PMLN)$	$\mathcal{O}\Big(\binom{P+2N_\mathrm{t}-1}{2N_\mathrm{t}-1}MLN\Big)$

the resulting asymptotic complexity. Firstly, it can be observed that, in the unreduced case, the complexities of the LS-based digital cancellers are asymptotically relative to the square of the total number of coefficients, meaning that they likely require more computations than the corresponding LMS-based cancellers. Namely, as also indicated in Table 4.1, the asymptotic complexities of the LMS-based solutions are either relative to the total number of coefficients or to the square of the number of static basis functions, the latter occurring if using a high-order nonlinear signal model. Hence, especially if a large number of memory taps is required for modeling the effective SI channel response, an LMS-based solution can be expected to be a more computationally efficient method for obtaining the SI channel estimate than using the LS algorithm. However, as mentioned earlier, the cost of this is the lower accuracy of the LMS estimate.

When employing the PCA-based model complexity reduction scheme, the asymptotic arithmetic complexities of the LS-based and LMS-based cancellers are in fact similar. This indicates that the complexity reduction scheme is better suited for the LS-based digital canceller. Investigating then the arithmetic complexities of the different signal models, it is evident that modeling the nonlinearity of the transmitter PA results in a higher computational complexity. Nevertheless, this can also be expected to improve the modeling accuracy and the cancellation performance, meaning that the higher computational complexity is likely a tolerable cost, especially when utilizing a highly nonlinear PA.

CHAPTER 5

Evaluating the Self-interference Cancellation Performance

This chapter provides simulation and measurement results where the proposed digital SI cancellation solutions are evaluated under different circumstances. Especially, in the measurements, the digital canceller complements an IBFD transceiver prototype, which suppresses the SI also in the analog/RF domain. This allows for evaluating the total SI cancellation performance, illustrating that true wireless IBFD operation is indeed possible. For reference, other IBFD prototype implementations, reported either earlier or parallel to the findings of this thesis, are also discussed comprehensively. The results and discussion within this chapter are based on the journal publications [**P1**–**P6**], as well as on the works in [**4, 17, 18, 82, 120**–**126, 133, 238**].

5.1 Demonstrator Implementations in Related Work

In the recent years, various IBFD transceiver prototypes or demonstrator implementations have been reported in the literature. Table 5.1 collects the key specifications and performance figures of the most notable prototype implementations, including also the ones discussed in, e.g., [**P4, P5**] that utilize the digital cancellers presented in this thesis. Note that, unless otherwise mentioned, the total amount of SI isolation listed in Table 5.1 for each prototype includes also the passive suppression, i.e., it is calculated as a difference between the transmit power and the residual SI power after all the cancellation stages. Further details regarding these prototypes are provided below, while the prototypes in [**P4, P5**] are discussed in Sections 5.3 and 5.4, respectively.

The pioneering work in [35] presents some of the very first experimental SI cancellation results, evaluating a 7×3 MIMO IBFD relay with separate TX and RX antennas. There, beamforming in the transmitter and linear digital cancellation in the receiver are utilized to suppress the SI signal over a bandwidth of 100 kHz. In total, almost 60 dB of cancellation is reported for the considered IBFD MIMO relay, even though the physical isolation between the TX and RX antennas is not included in this figure. Hence, the SI cancellation performance demonstrated in this early work is already rather promising,

Table 5.1: The key specifications and performance figures of the most notable IBFD prototype implementations. Note that all the prototypes are SISO transceivers, unless otherwise mentioned in the Structure-column

Prototype	Year	Frequency	Bandwidth	Structure	Analog canc.	Digital canc.	Total isolation
MIT I [35]	2007	370 MHz	0.1 MHz	7×3 MIMO relay	No	Yes	60 dB
Rice I [59, 63]	2010	2.4 GHz	0.625 MHz	Two antennas	Yes	Yes	80 dB
Stanford I [44]	2010	2.48 GHz	5 MHz	Three antennas	Yes	Yes	100 dB
Stanford II [103]	2011	2.4 GHz	10 MHz	Wired setup	Yes	Yes	73 dB
NYU Poly [117]	2012	915 MHz	7 MHz	Shared TX/RX antenna	Yes	No	59 dB
Rice II [61, 62]	2012	2.4 GHz	20 MHz	2×1 MISO	Yes	Yes	85 dB
Stanford III [28]	2013	2.45 GHz	80 MHz	Shared TX/RX antenna	Yes	Yes	110 dB
Rice III [73]	2014	2.4 GHz	20 MHz	Directional antennas	Yes	Yes	95 dB
Stanford IV [29]	2014	2.45 GHz	20 MHz	3×3 MIMO	Yes	Yes	104 dB
Chengdu [267]	2014	2.535 GHz	20 MHz	2×2 MIMO	Yes	Yes	114 dB
Bristol [140]	2015	1.89 GHz	20 MHz	Shared TX/RX antenna	Yes	No	83 dB
Fraunhofer [21]	2015	2.4 GHz	20 MHz	Shared TX/RX antenna	Yes	No	54 dB
Irvine [8]	2015	2.5 GHz	10 MHz	Multi-reconf. antenna	No	Yes	95 dB
Yonsei [50]	2015	2.52 GHz	20 MHz	Dual-polar. antenna	Yes	Yes	103 dB
Aalborg [212]	2016	2.4 GHz	80 MHz	Two antennas	Yes	Yes	100 dB
MIT II [119]	2016	2.45 GHz	20 MHz	Shared TX/RX antenna	Yes	No	78 dB
TCL [22]	2016	2.48 GHz	5 MHz	Two antennas	Yes	Yes	70 dB
TUT+Intel **[P4]**	2016	2.46 GHz	80 MHz	Shared TX/RX antenna	Yes	Yes	88 dB
TUT+Aalto **[P5]**	2017	2.56 GHz	80 MHz	Compact relay antenna	No	Yes	100 dB

likely contributing to the uprising of the IBFD technology as a potential method for improving the spectral efficiency of the future wireless systems.

Some time after this, the first IBFD prototype implementations utilizing both RF and digital cancellation were developed independently at Rice University [63] (reported in more detail later in [59]) and at Stanford University [44]. Even though the general architectures of these prototypes are rather similar, both the antenna structure and the RF cancellation techniques are fundamentally different in these two implementations. In particular, while the Rice prototype simply relies on the path loss between the TX and RX antennas to attenuate the SI, the Stanford prototype provides additional SI suppression in the antenna interface by adopting two TX antennas, one of them being half a wavelength further from the RX antenna than the other. Then, feeding the same transmit signal into the two antennas, destructive superposition will occur in the RX antenna, resulting in lower SI levels. The caveat of this technique is, however, that, even in theory, it only works for a single frequency, determined by the distance difference between the TX antennas and the RX antenna. For the considered bandwidth of 5 MHz, this limits the additional suppression provided by such an antenna cancellation technique in practice to 20 dB [44].

Another significant difference between [63] and [44] is the RF cancellation technique. Namely, in [63], the RF canceller utilizes an additional auxiliary TX chain to upconvert the cancellation signal from the baseband (see Fig. 2.5b), while in [44] the TX output signal is used to generate the RF cancellation signal (see Fig. 2.5a). As discussed in Section 2.3, the former approach has the benefit of additional flexibility since the cancellation signal is generated digitally, while the latter technique can suppress also all the impairments generated by the transmitter. In these early prototypes, the auxiliary TX–based method is reported to achieve higher cancellation levels, since the amount of RF cancellation in [63] is beyond 30 dB, while in [44] only 20 dB is obtained, albeit the former is reported for a significantly narrower bandwidth (625 kHz vs. 5 MHz). Including also the suppression provided by linear digital cancellation, the total amounts of SI cancellation are 80 dB and 100 dB in [63] and [44], respectively, while [59] reports 74 dB of overall SI cancellation for the same Rice prototype. The used transmit powers are reported to be from -5 dBm to 15 dBm in [63], and 0 dBm in [44]. Moreover, both of these prototypes are also shown to provide a higher throughput than a corresponding HD system. It should be noted, however, that the results reported in [59, 63] are only for a single subcarrier of 625 kHz, with the later works in [61, 62, 73] extending the Rice prototype to support several subcarriers and hence also wider bandwidths.

The next notable prototype is reported in [103], which presents the second revision of the Stanford prototype. It utilizes a similar RF cancellation architecture as the prototype in [44], while the antenna cancellation technique is not used anymore. Connecting the TX and RX together through a wire representing the coupling channel, roughly 40 dB of RF cancellation is achieved with the TX output–based technique, after which the SI is further suppressed by 30 dB using a linear digital canceller. Considering that the cancellation is performed here over a bandwidth of 10 MHz, obtaining 70 dB of SI suppression in total is a notable improvement over the earlier prototypes, albeit it was achieved only over a wire. The latter means that the SI channel is static and devoid of significant multipath components, making SI cancellation somewhat easier compared to a case where the SI signal is propagating wirelessly. However, it should

also be noted that, due to the wired setup, there is no physical isolation between the TX and RX chains, meaning that this prototype can potentially obtain a higher overall SI suppression performance when also passive isolation is introduced [103].

As opposed to the dual-antenna solutions used in [44, 59, 63], the prototype presented in [117] represents one of the earliest shared-antenna IBFD implementations. It utilizes a circulator, complemented with a balanced feed network, to provide 40–45 dB of passive isolation between the TX and RX chains when using an individual patch antenna for both transmission and reception. The amount of SI suppression is further increased by an active RF canceller, which utilizes the TX output signal and a VM to form the cancellation signal. Altogether, 59 dB of SI cancellation is reported over a 7-MHz bandwidth when using a transmit power of 0 dBm.

Improved versions of Rice University's initial prototype are then reported in [61, 62, 73]. The prototype demonstrated in [61, 62] is a multiple-input and single-output (MISO) IBFD transceiver with two TX antennas and one RX antenna, which are positioned around a device to provide a higher amount of passive isolation, measured to be between 60 and 70 dB. The received SI signal is then first cancelled by an active auxiliary TX chain–based RF canceller, which basically extends the single-subcarrier solution of [63] to support several adjacent subcarriers. Moreover, as the prototype has two transmitters, the RF cancellation signal is generated using both of the transmit signals. This solution obtains roughly 15–20 dB of active RF cancellation over the considered bandwidth of 20 MHz. When complemented with linear digital cancellation, the overall amount of SI suppression for this prototype is 85 dB with a total transmit power of 5 dBm. The prototype reported in [73], on the other hand, utilizes directional antennas with cross-polarization and absorptive shielding to provide 70 dB of physical isolation between the TX and RX chains. The SI is then further cancelled by a similar active auxiliary TX chain–based RF canceller as in [61, 62], although now modified for a SISO architecture. After applying also linear digital cancellation, the overall amount of SI suppression over 20 MHz is in the order of 95 dB with this solution, using a transmit power of 7 dBm. Both of the proposed IBFD prototypes in [61, 62, 73] are also shown to obtain higher data rates than comparable HD implementations.

The third prototype implementation of the Stanford group, reported in [28], represents in many respects the state of the art, even though it is already four years old. There, a shared-antenna architecture is adopted, the TX and RX chains being isolated by a circulator. The RF canceller utilizes again the TX output, dividing it into 16 parallel replicas with different fixed delays. The actual RF cancellation signal is then constructed by attenuating each of the replicas by a specific amount, determined by a control algorithm, after which they are combined and summed up with the total received signal. The final cancellation stage in the digital domain utilizes then a PH signal model to reconstruct also the nonlinear distortion within the residual SI. Hence, [28] reports one of the first measurement-based evaluations of a nonlinear digital canceller. The overall amount of SI cancellation with a 20-dBm transmit power and an instantaneous bandwidth of 80 MHz is reported to be 110 dB, consisting of 62 dB of analog SI suppression and 48 dB of digital cancellation. This is sufficient to suppress the SI to the level of the receiver noise floor, and consequently this prototype is also demonstrated to nearly double the throughput when compared to a HD system. Moreover, using a similar architecture, the same amount of cancellation is also reported for a transmit

power of 25 dBm and a bandwidth of 20 MHz. All in all, the cancellation performance reported in [28] still represents the benchmark for any IBFD transceiver.

The authors of [28] have also extended their SI cancellation architecture to support a 3×3 MIMO transceiver [29]. The proposed IBFD MIMO prototype consists of three TX–RX pairs where each pair shares a common antenna and the passive isolation is achieved using a circulator. The RX chains are preceded by a 56-tap 3×3 RF canceller, followed by a nonlinear digital canceller after the analog-to-digital conversion in each receiver. Using a sum transmit power of 20 dBm (divided equally between the three transmitters), the SI is cancelled in total by roughly 104 dB. This is sufficient to cancel the residual SI practically to the level of the receiver noise floor in each RX chain.

While the prototype implementations in [28, 29, 35, 44, 61–63, 73, 103, 117] precede the work reported in this thesis and can therefore be considered prior art, the other IBFD prototypes listed in Table 5.1 have been developed either in parallel to or after it. Of these, the earliest is the prototype in [267], where a 2×2 IBFD MIMO transceiver using separate TX and RX antennas is reported and evaluated. It has two 4-tap RF cancellers for both receivers, utilizing the output signals of the two transmitters to generate the RF cancellation signal. Further SI suppression is then performed by the digital canceller, which also uses the TX output signals, downconverted and digitized, to regenerate the residual SI. This facilitates the use of linear processing in the digital domain while still suppressing the TX-induced impairments. Due to the rather large spatial separation between the TX and RX antennas, the physical isolation of this prototype is nearly 40 dB, meaning that, with a 23-dBm transmit power, the total SI power before RF cancellation is -15 dBm. Then, after 43 dB of RF cancellation and 33 dB of digital cancellation, the SI is cancelled to the level of the receiver noise floor, the overall amount of SI cancellation being 114 dB over 20 MHz. The authors also report that the implemented IBFD prototype can obtain a data rate two times higher than that of a corresponding HD system.

As opposed to this, the IBFD prototype proposed in [140] utilizes an EBD to separate the TX and RX chains when using a shared TX/RX antenna. As already discussed in Section 2.3, an EBD provides the isolation by mimicking the impedance of the antenna, which results in an exact copy of the transmit signal that is reflected to the RX chain [56, 140, 141, 263]. Using then a hybrid transformer to phase shift this copy by 180° and combining it with the overall RX signal, the SI signal coupling to the receiver is significantly attenuated. With a center frequency of 1.89 GHz and a transmit power of 10 dBm, the EBD is shown to suppress the SI by 45 dB over a bandwidth of 20 MHz. This is then complemented by an auxiliary TX chain–based RF canceller, which provides further 38 dB of SI suppression, the total amount of cancellation being 83 dB.

The work in [21] reports results for a similar shared-antenna implementation with an auxiliary TX chain–based RF canceller, although in this prototype a circulator is used to provide the passive isolation between the TX and RX chains. Furthermore, to improve the cancellation performance, the I/Q imbalance of each TX and RX chain is also compensated for by estimating the necessary coefficients, which is done during a separate training period. After this, the estimated coefficients are used to pre-compensate for the I/Q imbalance, allowing for accurate estimation of the actual SI coupling channel. With this procedure, the SI is cancelled in total by 54 dB over a bandwidth of 20 MHz, of which 39 dB is provided by the RF canceller.

In [8], an IBFD prototype utilizing a multi-reconfigurable antenna (MRA) is proposed and evaluated. In particular, an MRA is a dynamically reconfigurable antenna whose properties in terms of frequency, radiation pattern, and polarization can easily be modified by choosing the appropriate input configuration [146]. In this case, the radiation pattern of the MRA structure is configured such that the received SI power is minimized, and thereby a high amount of passive isolation between the TX and RX chains is obtained. In this prototype, such a scheme provides nearly 65 dB of passive isolation over the considered bandwidth of 10 MHz. When complemented with linear cancellation in the digital domain, the overall SI suppression of this prototype is 95 dB with a transmit power of 5 dBm. Moreover, nearly a twofold improvement in the throughput in comparison to a corresponding HD device is reported.

Considering then the prototype in [50], there the passive TX–RX isolation is improved by utilizing a dual-polarized antenna while also performing active RF and digital cancellation to further suppress the SI. In particular, a dual-polarized antenna transmits and receives signals of different polarization, resulting in a greatly attenuated leakage from the TX port to the RX port [56]. If the TX and RX polarizations were completely orthogonal, the isolation would be perfect, while only 40 dB of SI suppression is achieved at this stage in [50]. A TX output–based RF canceller is then used to further suppress the SI by 18 dB, after which the resulting signal is digitized. Then, the SI is digitally cancelled in the frequency domain using a real-time field-programmable gate array (FPGA) implementation and a linear signal model, providing 43 dB of additional SI suppression. Overall, the proposed architecture in [50] is capable of canceling the SI by 103 dB over 20 MHz. In addition, the IBFD prototype is also shown to obtain a 90% increase in the throughput compared to a corresponding FDD implementation.

Another dual-antenna IBFD transceiver prototype with active RF and digital cancellation stages is reported in [212]. This prototype has a rather large antenna separation, resulting in 50 dB of physical isolation, the rest of the SI being then suppressed by the active cancellers. The RF-domain canceller utilizes an auxiliary TX chain–based solution, where a digitally generated cancellation signal is upconverted and subtracted from the RX input signal. The cancellation in the digital domain, on the other hand, is performed with one of two options: a nonlinear canceller that is using the original transmit data and a Hammerstein PA model to regenerate the nonlinearly distorted SI, or by downconverting and digitizing the TX output signal and using it to generate the digital cancellation signal. Altogether, the SI is suppressed by 100 dB using a transmit power of 20 dBm and an instantaneous bandwidth of 80 MHz.

In [119], measurement results are presented for a prototype consisting of a single high-isolation antenna and a multi-tap RF canceller. The latter utilizes the TX output signal to form the cancellation signal, using four taps with different delays and with separate phase and amplitude control. The correct phase and amplitude values are estimated using a modified version of the dithered linear search (DLS) algorithm, which requires knowledge of only the residual power after RF cancellation for parameter learning, resulting in a very simple learning rule. In a realistic indoor environment, the prototype in [119] obtains 78 dB of SI cancellation in total when using a 30-dBm transmit power and a 20-MHz bandwidth, of which 22 dB is provided by the four-tap RF canceller. When increasing the bandwidth to 120 MHz, the implemented RF canceller is capable of canceling the SI by 13.3 dB, resulting in an overall SI suppression of 70 dB.

The work in [22] presents yet another different approach for implementing an IBFD device. Namely, there an auxiliary TX chain–based RF canceller is used to regenerate and cancel the SI in the RF domain, but such that both the actual and auxiliary TX chains are linearized by means of digital predistortion. This means that the PA-induced nonlinear distortion can be omitted in the digital cancellation stage. With separate TX and RX antennas placed 30 cm apart, the prototype is shown to suppress the SI by 70 dB in total when transmitting a 5-MHz signal with a 20-dBm average transmit power. Of this, the amount of RF cancellation is roughly 40 dB, while the digital canceller, utilizing the signal model presented in [26] that incorporates the effects of DAC nonlinearities and I/Q imbalance, suppresses the SI by 11 dB.

The above prototypes represent the current state of the art in the literature, and thereby constitute the proper context for the SI cancellation performance achieved using the digital cancellation solutions presented in this thesis and reported in **[P4, P5]**. Table 5.1 presents also the key performance figures of these prototype implementations, while further details are provided below in Sections 5.3 and 5.4.

5.2 Simulated Self-interference Cancellation Performance

Let us first verify the different digital cancellation solutions with a realistic waveform simulator that models all the significant analog impairments within a 2×2 MIMO IBFD transceiver, reported in detail in **[P6]**. The waveform simulator is implemented in Matlab and it follows the architecture in Fig. 2.5a, extended to a 2×2 MIMO scenario, to produce a realistic residual SI signal for the digital canceller. The used transmit signals are 20 MHz orthogonal frequency-division multiplexing (OFDM) waveforms following the specifications of LTE DL signals, the transmit power being defined as the combined power of these individual transmit signals after the PAs. Moreover, the same system parameters that were used in Section 3.2, listed in Tables 3.1 and 3.2, are used also in the waveform simulations, complemented with the additional parameters listed in Table 5.2. The only exception is the phase noise model, whose characteristics are now as shown in Fig. 5.1. The adopted phase noise model corresponds to a real-life charge-pump-type PLL–based oscillator to ensure as realistic results as possible [165]. Furthermore, no received signal of interest is present in the simulations to focus on the overall SI cancellation performance.

The MIMO SI coupling channel used in the simulations is randomly generated for each realization, although it is always set to have the desired K value, adopted from the measurement results reported in [59]. Similarly, the RF cancellation signals are generated using a random error component, and consequently the amount of RF cancellation varies from one realization to the next. Nevertheless, the average amount of RF cancellation is as specified in Table 3.1. The forthcoming results are generated by running 20 independent realizations for a given set of parameter values and then measuring the average residual SI power of these realizations. The only exception are the power spectral densities (PSDs), which represent only the outcome of a typical realization within an individual receiver for illustrative purposes. Hence, no major conclusions should be drawn from the PSDs alone, as they include no averaging, although they still represent a rather accurate scenario in terms of the true cancellation performance.

Table 5.2: The additional default parameters used in the waveform simulations.

Parameter	Value
Number of TX/RX chains (N_t/N_r)	2/2
Sampling frequency	122.88 MHz
Level of TX crosstalk before the PAs	−15 dB
Level of TX crosstalk after the PAs	−15 dB
Transmit waveform	OFDM
SI channel length	20 taps
SI channel K value	35 dB
Parameter estimation sample size (N)	30 000
Nonlinearity order of the cancellers (P)	5
Number of pre-cursor taps (M_1)	10
Number of post-cursor taps (M_2)	20

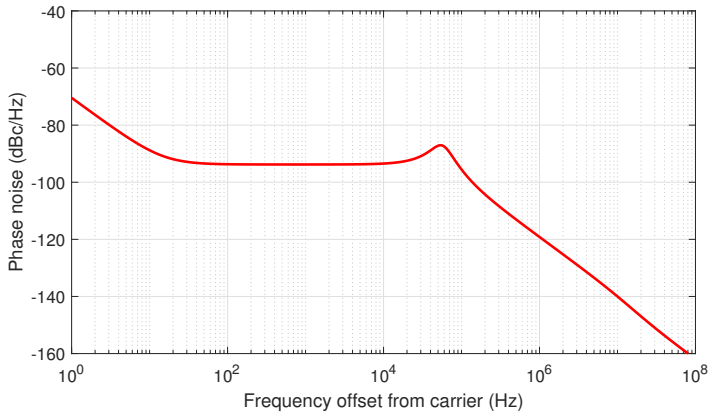


Figure 5.1: The phase noise characteristics used in the waveform simulator, taken from [165].

Using the waveform simulator, the following digital cancellation solutions are evaluated:

- digital cancellation using the linear signal model, presented in Section 4.2.1;
- digital cancellation using the widely linear signal model, presented in Section 4.2.2;
- digital cancellation using the nonlinear signal model, presented in Section 4.2.3;
- digital cancellation using the nonlinear signal model incorporating crosstalk and I/Q imbalance, presented in Section 4.2.4, without model complexity reduction;
- digital cancellation using the nonlinear signal model incorporating crosstalk and I/Q imbalance, presented in Section 4.2.4, with the model complexity reduction scheme presented in Section 4.3.3.

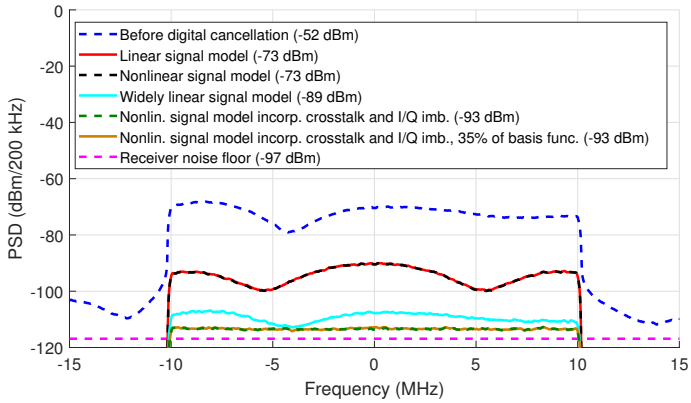


Figure 5.2: The PSDs of the signal after the different SI cancellation stages in the waveform simulator. Here, the RX input is used as the reference point for all the signal powers, meaning that they do not include the RX gain.

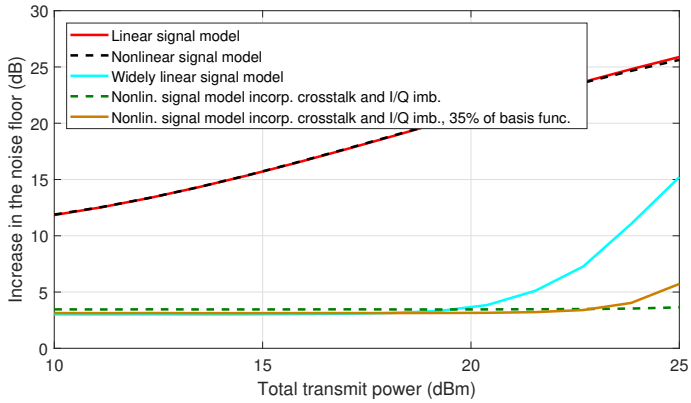


Figure 5.3: The residual SI–induced increase in the noise floor with respect to the total transmit power.

In all the cases, the unknown signal model coefficients are learned using the LS-based parameter estimation scheme presented in Section 4.3.1. The estimation is done over the specified amount of samples (N), while the cancellation performance itself is evaluated for a different set of signal samples. Note that in this case oversampling is not necessary when generating the nonlinear basis functions as the initial sampling frequency is high enough to capture also the higher-order nonlinearities.

First, Fig. 5.2 shows the PSDs of the signal after the different SI cancellation stages using the default system parameters. It confirms the observation made in Section 3.2 regarding the dominant nature of the I/Q imbalance, as neither the nonlinear nor the linear canceller perform very well. On the other hand, the digital canceller utilizing the widely linear signal model obtains significantly higher levels of SI cancellation, while the case where both the I/Q imbalance and the nonlinear distortion are modeled results in the lowest residual SI power, as can be expected. The 4 dB increase in the noise floor

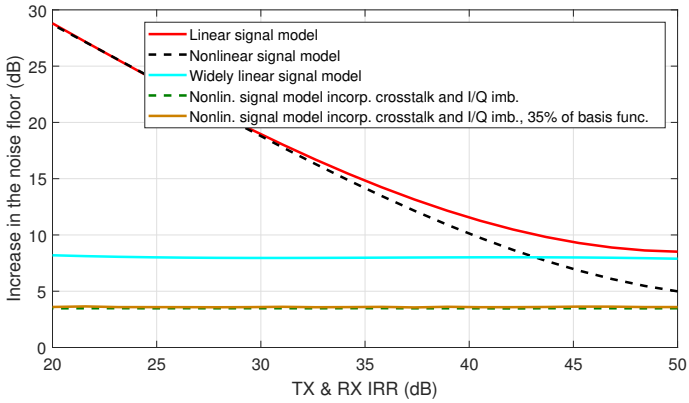


Figure 5.4: The residual SI–induced increase in the noise floor with respect to the IRR of the TX and RX chains.

even with the best digital cancellers can be attributed to the RF canceller output noise, as concluded already in Section 3.2. It can also be observed that retaining only 35% of the basis functions of the nonlinear signal model incorporating crosstalk and I/Q imbalance still results in the same cancellation performance. This illustrates the benefits of the proposed PCA-based model complexity reduction scheme.

To investigate the effect of the transmit power in more detail, Fig. 5.3 shows the increase in the noise floor, caused by the IBFD operation, with respect to the total transmit power. Again, in accordance with the conclusions made in Section 3.2, the signal models omitting the I/Q imbalance perform rather poorly, even with the lowest transmit powers. As for the widely linear signal model, it is capable of accurately modeling the residual SI up to a transmit power of approximately 20 dBm, after which the PA-induced nonlinearities start to become a significant factor. Beyond this point, modeling of both the I/Q imbalance and the nonlinearities is required to efficiently suppress the residual SI. The exact transmit power level where this transition occurs depends on the characteristics of the TX PA; if the PA is highly nonlinear, modeling of the nonlinear distortion is required with lower transmit powers, whereas the opposite is true for a more linear PA. Also note that the PCA-based model complexity reduction scheme increases the cancellation performance with the lower transmit powers as the smaller number of coefficients improves the estimation accuracy. With the highest transmit powers, however, more than 35% of the basis functions are needed to achieve the full modeling accuracy, as can be observed in Fig. 5.3.

Considering then the effect of I/Q imbalance, Fig. 5.4 shows the increase in the noise floor with respect to the IRR of the TX/RX chains, assuming that both have the same IRR. As can be expected, with the higher IRRs, the nonlinear signal model is more accurate than the widely linear model as then the PA nonlinearity is the dominant form of distortion. However, if the IRR is less than 43 dB, the digital canceller utilizing the widely linear signal model is the better option of these two. Moreover, as can be expected, the signal model considering both the I/Q imbalance and the nonlinear distortion retains high accuracy regardless of the IRR value, the cancellation performance of the reduced version being again nearly identical to the one utilizing all the basis functions.

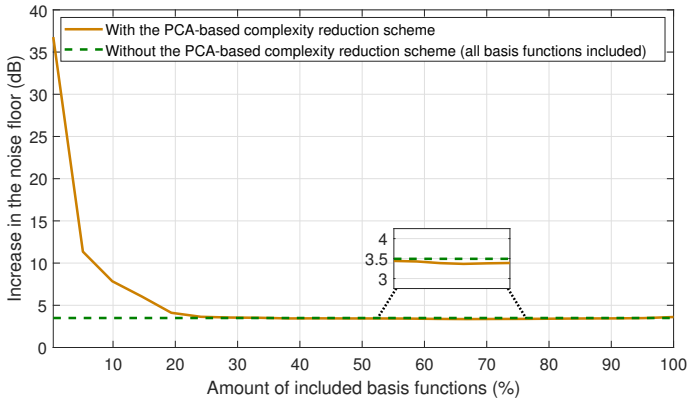


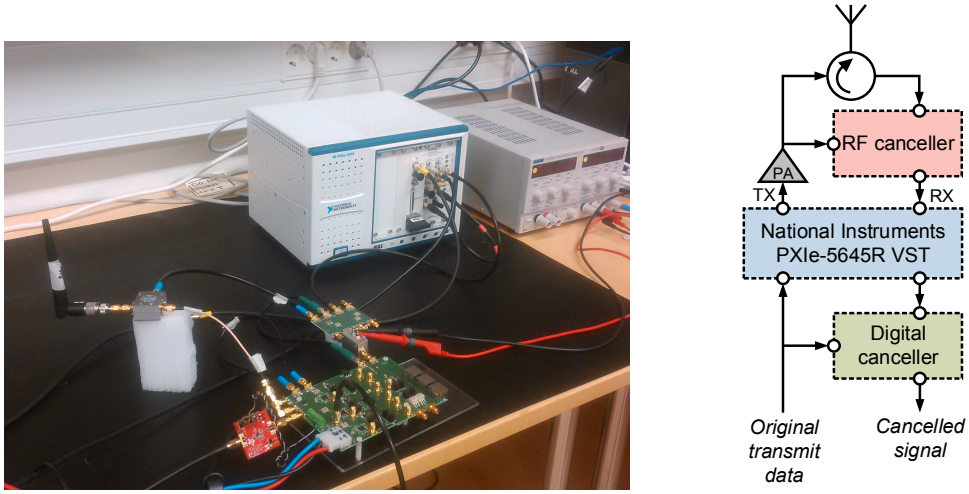
Figure 5.5: The residual SI–induced increase in the noise floor with respect to the amount of included basis functions.

To gain more insight into the model complexity reduction scheme, Fig. 5.5 shows the increase in the residual noise floor with respect to the relative number of basis functions retained after the PCA processing. With the chosen example system parameters, retaining only 30% of the basis functions results still in roughly the same cancellation performance as using all of them. In fact, as can be observed also from this figure, reducing the number of terms slightly increases the amount of SI cancellation as then less coefficients must be estimated. This results in the improved estimation accuracy of the remaining coefficients, as already discussed. Hence, these observations indicate that it is highly beneficial to carefully control the number of basis functions in the more complex signal models.

Overall, the above findings show that linear SI cancellation alone is insufficient for a realistic IBFD transceiver with reasonable RF component quality. Depending on the level of I/Q imbalance, either nonlinear or widely linear modeling is required in the digital canceller to obtain sufficient SI regeneration accuracy. Furthermore, if both of these impairments are present in the system, a signal model incorporating both widely linear and nonlinear modeling is required in the digital canceller. Below, Sections 5.3 and 5.4 confirm the need for nonlinear modeling using real-life IBFD prototype implementations.

5.3 Measured Self-interference Cancellation Performance of a Generic Inband Full-Duplex Device

Then, in order to demonstrate the real-life performance of the digital cancellation solutions, they are incorporated into an IBFD prototype implementation, designed to be a generic SISO radio device. The hereby obtained results are reported also in **[P4]** and partially in **[125]**, the latter providing some further technical details. The actual measurement setup is as shown in Fig. 5.6a, while Fig. 5.6b illustrates the basic structure of the prototype. Moreover, the relevant parameters of the measurement setup and the used digital cancellers are listed in Table 5.3. Similar to the simulations, no received signal of interest is present in the forthcoming measurement results to concentrate



(a) **(b)**

Figure 5.6: (a) The measurement setup used in evaluating the digital cancellation performance of the generic IBFD transceiver prototype, and (b) the basic structure of the evaluated prototype.

Table 5.3: The essential parameters of the measured generic IBFD transceiver prototype.

Parameter	**Value**
Center frequency	2.46 GHz
Transmit waveform	OFDM
Bandwidth	20/40/80 MHz
Transmit power	6–8 dBm
TX/RX sampling rate	120 MHz
PA gain	24 dB
Parameter estimation sample size (N)	500 000
Number of pre-cursor taps (M_1)	10
Number of post-cursor taps (M_2)	20
Order of the nonlinear canceller (P)	11

on evaluating the actual SI cancellation performance. In this implementation, the National Instruments (NI) PXIe-5645R vector signal transceiver (VST) is used both as the transmitter and the receiver, the TX output signal being further amplified by a low-cost PA (Texas Instruments CC2595 [242]). The PA output is then connected to the antenna via a circulator, which allows the usage of a single antenna while still isolating the TX and RX chains to some extent, as discussed earlier.

After the circulator, the received signal is fed to the first active SI cancellation stage, a multi-tap RF canceller. It is discussed in detail in **[238]**, and also illustrated on a general level in Fig. 2.5a. This particular RF canceller has three taps, whose phases and amplitudes are tuned with a Hittite HMC631LP3 VM to minimize the SI power in the canceller output [89]. The correct phase and amplitude values of the individual taps are

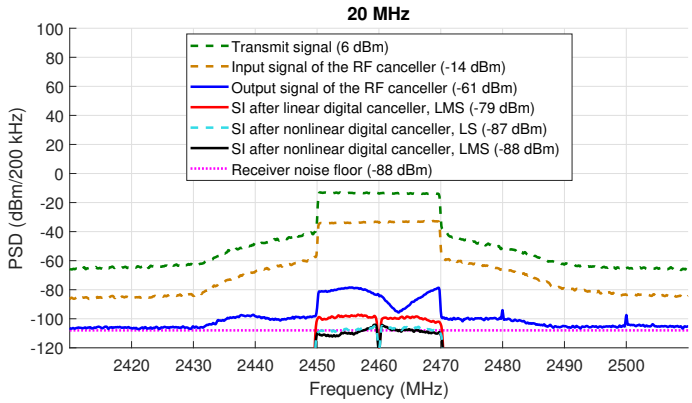
learned with a basic digital LMS algorithm running on a separate FPGA, meaning that the tap signals and the RF canceller output signal must also be digitized for parameter learning. Having a digital parameter learning algorithm for the RF canceller simplifies its RF design while also allowing for more flexibility when experimenting with different control algorithms [**99, 238**].

Having performed RF cancellation, the signal is fed to the RX input of the VST and digitized for further offline post-processing.[1] Digital cancellation is then performed on the digitized and recorded signal, using Matlab. Here, only the cancellation performance of the linear and nonlinear digital cancellers is demonstrated as the PA-induced nonlinear distortion is the dominant impairment in this prototype implementation, owing to the low-cost PA. I/Q imbalance is not an issue due to the high IRR of the PXIe-5645R VST. Moreover, in the nonlinear canceller, the basis functions are generated with an incremented sampling rate of 240 MHz to avoid excessive aliasing of the nonlinear terms, after which they are decimated back to the RX sampling rate of 120 MHz for parameter learning and cancellation. In the linear canceller, no such oversampling is required.

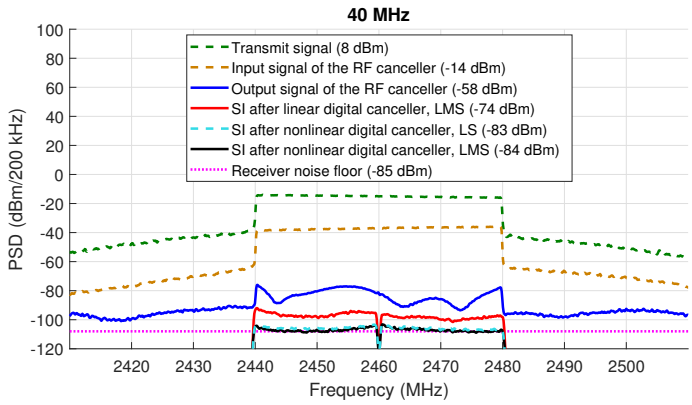
As for the parameter estimation within the digital canceller, both the LMS-based and LS-based solutions are utilized in the forthcoming measurement results, although the LMS is still considered the primary option and the performance of the LS-based solution is only shown in the results for reference. In particular, as discussed in Section 4.3.4, the LMS-based parameter learning and cancellation is computationally less demanding than using the LS to estimate the coefficients, and consequently the digital cancellation performance obtained with the LMS represents a more realistic scenario also in terms of commercial applications. What is more, under rapidly changing SI channel conditions, the adaptive LMS algorithm also facilitates the tracking of the channel. To ensure that the shown results correspond to the highest achievable cancellation performance, the LMS is iterated over two epochs of the specified amount of data, and the performance is then measured from the last epoch. For clarity and brevity of presentation, the residual SI after the LS parameter estimation is only shown for the nonlinear signal model, using the same amount of data as for the LMS learning.

In order to evaluate the true digital SI cancellation performance, the noise in the forthcoming measurement results is reduced by averaging the cancelled signal over many successive repetitions of the same transmit signal. Namely, as the residual SI is a function of the original transmit signal, it is not affected by the averaging, while the noise is reduced as it is obviously different from one repetition to the next. The SI coupling channel can also be expected to be identical between the different repetitions, as they occur within a very short time span. Moreover, since the LMS algorithm is allowed to converge to the steady-state before measuring the cancellation performance and performing the averaging, also the coefficient estimates can be expected to be static over the successive repetitions. This is obviously the case also for the LS parameter estimation, as it implicitly assumes the same coefficients for the whole estimation period. Thus, the residual SI is of the same form for all the repetitions, and consequently the averaging only removes the noise. All in all, when interpreting the forthcoming results, it should therefore be taken into account that the total noise-plus-interference floor is obtained by summing the noise into the shown residual SI.

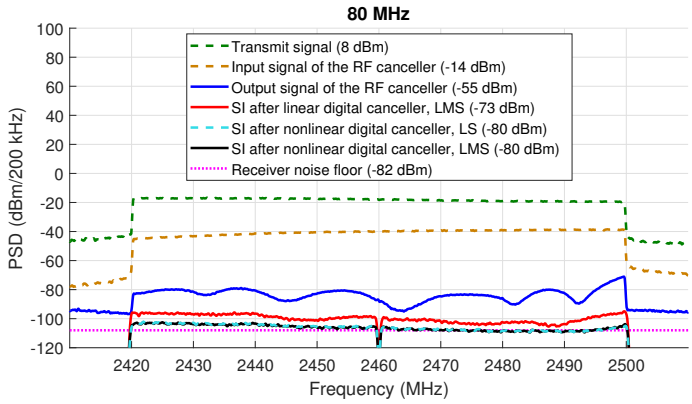
[1] Even though the results reported in this thesis are generated using offline processing, the Author has also reported a real-time implementation of the same nonlinear digital canceller [**120**].



(a)



(b)



(c)

Figure 5.7: The PSDs after the different SI cancellation stages in the generic IBFD device for (a) 20-MHz bandwidth, (b) 40-MHz bandwidth, and (c) 80-MHz bandwidth.

Figure 5.7 then shows the PSDs of the SI signal at different interfaces of the IBFD transceiver for the three considered bandwidths, the signal powers being referred to the RX input. Investigating first the 20-MHz case in Fig. 5.7a, here the residual SI power after RF cancellation is −61 dBm, which is clearly above the receiver noise floor and hence calls for further cancellation in the digital domain. It can firstly be observed that, due to the highly nonlinear low-cost PA, the linear digital canceller is not capable of fully suppressing the residual SI, achieving roughly 18 dB of cancellation. On the other hand, the nonlinear canceller manages to suppress the SI to the level of the receiver noise floor, canceling it by over 25 dB with both the LMS-based and LS-based parameter learning solutions. Hence, in total, the amount of obtained SI cancellation is 94 dB, of which 20 dB is physical isolation provided by the antenna and the circulator, while 47 dB of the overall suppression is contributed by the RF canceller. The small difference in the residual SI powers after the LMS-based and LS-based cancellers can likely be attributed to some very slight temporal changes in the SI coupling channel, which the LS algorithm is incapable of tracking. When including also the noise, the residual power after the nonlinear digital canceller is −80 dBm, indicating that this particular RF canceller prototype produces a rather significant amount of noise. Hence, even though the SI cancellation performance itself is very high, both in the RF and digital domains, a less noisy RF canceller is required to manufacture a fully operational shared-antenna IBFD transceiver with no excessive SINR loss.

The nonlinear digital canceller copes well also with the wider bandwidths, since the residual SI is still cancelled close to the receiver noise floor in the 40-MHz and 80-MHz cases, as can be observed from Figs. 5.7b and 5.7c. Moreover, similar to the 20-MHz scenario, the residual SI after linear digital cancellation is roughly 10 dB above the noise floor also with these wider bandwidths, indicating that nonlinear modeling is indeed necessary. With a bandwidth of 40 MHz, the total amount of SI cancellation is therefore 92 dB, of which 44 dB is provided by the RF canceller, while the nonlinear digital canceller suppresses the SI by 26 dB. Correspondingly, in the 80-MHz scenario, the SI is suppressed by 88 dB in total, consisting of 41 dB of RF cancellation and 25 dB of nonlinear digital cancellation. Thus, the overall SI cancellation performance remains on a high level even when using such wide instantaneous bandwidths. Moreover, the chosen parameter estimation scheme does not seem to affect the cancellation accuracy, indicating that the LMS-based solution is capable of obtaining sufficiently accurate coefficient estimates, regardless of its lower computational complexity in comparison to LS estimation. When including also the noise, the overall residual power after digital cancellation is −76 dBm for the bandwidth of 40 MHz and −75 dBm for the bandwidth of 80 MHz.

In conclusion, with the considered transmit powers of 6–8 dBm and instantaneous bandwidths of 20–80 MHz, the implemented IBFD prototype is capable of canceling the residual SI to the level of the receiver noise floor when using the developed nonlinear digital canceller with LMS-based parameter learning. Even though the utilized RF canceller is producing rather significant levels of noise into the digital canceller input signal, the performance of the digital canceller is still sufficient to efficiently suppress the residual SI. Hence, by further improving the RF canceller design, a fully functional IBFD transceiver can be implemented by utilizing the proposed nonlinear digital canceller in conjunction with the analog SI suppression techniques.

5.4 Measured Self-interference Cancellation Performance of an Inband Full-Duplex Relay

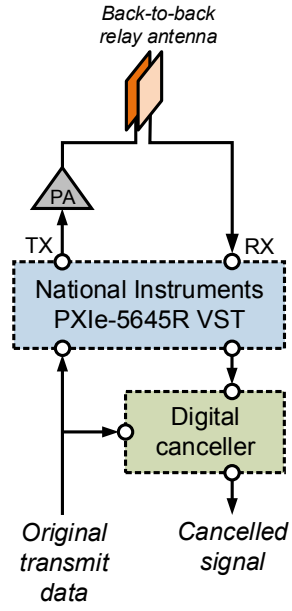
Another IBFD prototype implementation utilizing the digital cancellers proposed in this thesis is reported in **[P5]**, where a relay-type scenario is considered. Namely, in this prototype, a compact high-isolation back-to-back relay antenna, designed for a center frequency of 2.56 GHz, is used as the radiating element. The antenna element is designed such that the transmit and receive directions are on the opposite sides of the structure, which provides a high level of inherent physical isolation, especially when complemented with so-called wavetraps that further decrease the SI coupling between the TX and RX sides [88].

The high physical isolation of the antenna structure also means that no active RF canceller is actually needed in the prototype. In particular, the power of the SI signal coupling to the receiver is suppressed sufficiently low already by the antenna structure, meaning that it is within the dynamic range of the RX chain without any additional analog cancellation. Hence, the residual SI can be suppressed by performing active cancellation only in the digital domain. This obviously decreases the complexity of the overall SI cancellation architecture as no additional RF hardware is required. However, the burden on the digital canceller is correspondingly greater, as it must reduce the SI level by a somewhat larger amount.

The measurement setup is now as shown in Fig. 5.8a, while the general structure of the prototype is illustrated in Fig. 5.8b. Moreover, all the relevant parameters of the measurement setup and the digital cancellers are listed in Table 5.4. The used measurement location is in the Kampusareena library, at Tampere University of Technology, meaning that it represents a realistic indoor deployment scenario with various reflecting surfaces. Again, the NI PXIe-5645R VST is used both as the transmitter and the receiver, the TX output signal being also amplified by an external PA (Mini-Circuits ZVE-8G+ [168]).

Now, the PA output signal is directly fed to the TX port of the high-isolation antenna, the SI coupling to the RX port via the surrounding reflections and the direct leakage through or around the structure (although the latter is greatly weakened by the wavetraps). As there is no RF canceller in this prototype, the RX port of the antenna structure is then directly connected to the receiver (PXIe-5645R VST) where the received signal is digitized for offline post-processing. Similar to the results in Section 5.3, the cancellation performance of both the linear and nonlinear digital cancellers is evaluated when receiving only the SI from the own transmitter, i.e., no signal of interest is present. Moreover, in the nonlinear digital canceller, the basis functions are again generated with a sampling frequency of 240 MHz to avoid excessive aliasing of the nonlinear terms. Also note that now the digital cancellers must use a larger number of taps than in the generic IBFD transceiver where also RF cancellation is performed. This stems from the higher SI power entering the digital domain, whose efficient cancellation requires more accurate modeling.

The parameter estimation is carried out with both the LMS-based and LS-based solutions, the latter being again applied only to the nonlinear signal model. Both parameter estimation schemes use the specified amount of data for learning, the LMS algorithm being iterated now over three epochs to ensure full convergence. The SI

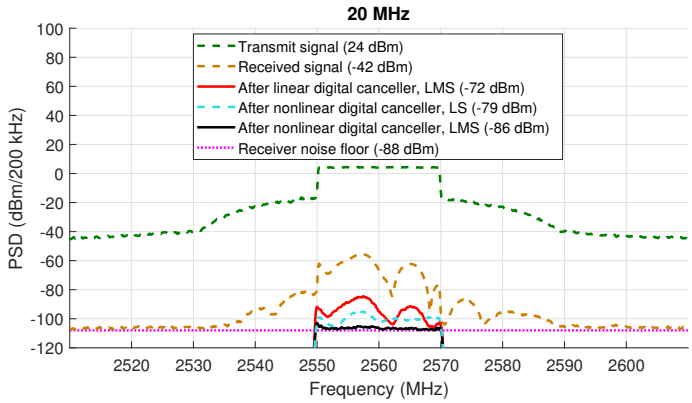


(a) **(b)**

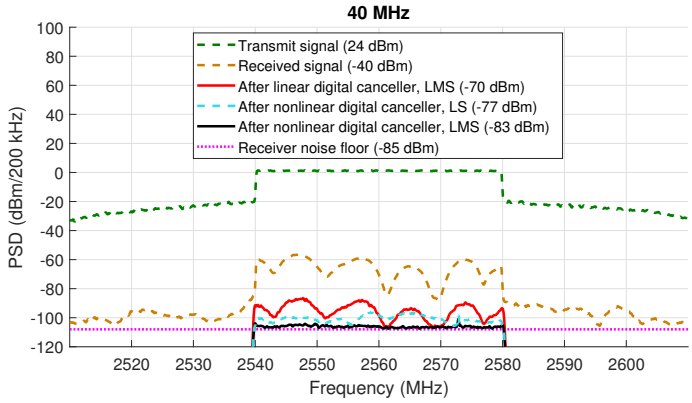
Figure 5.8: (a) The measurement setup used in evaluating the digital cancellation performance of the IBFD relay prototype, and (b) the basic structure of the relay prototype.

Table 5.4: The essential parameters of the measured IBFD relay prototype.

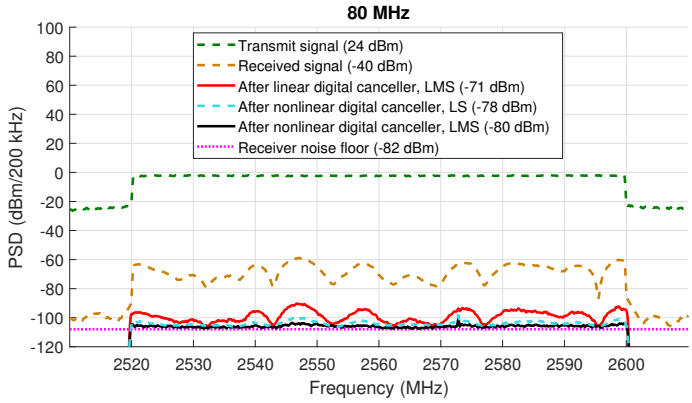
Parameter	**Value**
Center frequency	2.56 GHz
Transmit waveform	OFDM
Bandwidth	20/40/80 MHz
Transmit power	24 dBm
TX/RX sampling rate	120 MHz
PA gain	36 dB
RX losses	4 dB
Parameter estimation sample size (N)	1 000 000
Number of pre-cursor taps (M_1)	25
Number of post-cursor taps (M_2)	50
Order of the nonlinear canceller (P)	11



(a)



(b)



(c)

Figure 5.9: The PSDs after the different SI cancellation stages in the IBFD relay for (a) 20-MHz bandwidth, (b) 40-MHz bandwidth, and (c) 80-MHz bandwidth.

cancellation performance of the LMS-based solution is then measured over the last epoch. However, as opposed to the results reported in Section 5.3, now no additional noise is produced due to the lack of RF canceller. Hence, the digital SI cancellation performance can be evaluated reliably without any noise removal procedure. For this reason, the forthcoming results show the overall residual signal power, including also the noise.

The PSDs of the signal after the different SI cancellation stages are shown in Fig. 5.9 for all the considered bandwidths, using again the RX input as the reference point for the different power levels. Investigating first the isolation provided by the antenna, it can be observed that it is in the order of 60 dB for each bandwidth when considering also the cable losses in the RX path (4 dB). Hence, the power of the residual SI signal entering the digital domain is roughly -40 dBm in all the cases, which is already sufficiently low for the RX chain to operate without saturation or excessive distortion. Such a power level is also within the dynamic range of the ADC, meaning that the quantization noise remains well below the receiver noise floor. Consequently, no active cancellation is required before the ADC.

Investigating then the digital cancellation performance, it is clear that also now the linear digital canceller is incapable of fully suppressing the residual SI. Namely, the residual SI power after linear cancellation is still 10–15 dB above the receiver noise floor, rendering the IBFD operation infeasible. On the other hand, the nonlinear canceller with LMS-based parameter learning is capable of very efficient SI cancellation with all the considered bandwidths, reducing the overall noise-plus-interference power practically to the level of the receiver noise floor. Hence, the amount of digital SI cancellation is beyond 40 dB for each considered bandwidth and, taking into account the RX cable losses, the overall amounts of SI suppression are 106 dB, 103 dB, and 100 dB over 20 MHz, 40 MHz, and 80 MHz, respectively.

However, as opposed to the results in Section 5.3, now the accuracy of the LS parameter estimates is observably lower than those obtained with the LMS algorithm. This likely stems from the more time-variant nature of the measurement environment, resulting in the SI channel changing even during the rather short recorded signal sequence. Consequently, tracking of the SI channel coefficients is required, and for this reason the LMS is better suited for the parameter estimation in this case, thanks to its adaptive nature.

In general, these measurement results obtained with an IBFD relay prototype demonstrate the flexibility of the developed LMS-based nonlinear digital cancellation algorithm, as it is shown to be capable of efficient SI cancellation even without any RF cancellation stage. This greatly reduces the overall complexity of the IBFD device as only one active cancellation stage is required, albeit the digital canceller requires now more memory taps due to the higher cancellation demands. Nevertheless, no additional RF hardware is required, which likely results in a lower overall cost for the transceiver.

CHAPTER 6

Applying Inband Full-Duplex Communications on a System Level: Self-backhauling Access Node

This chapter investigates and analyzes the possibility of wireless self-backhauling with an IBFD-capable AN, which is also serving DL and UL UEs simultaneously on the same frequency band. Such an IBFD solution is compared to two reference schemes, which divide at least some of the communications tasks into separate time slots. The system is then analyzed in terms of transmit power minimization under given QoS conditions, all of the results being derived in closed form. The contents of this chapter are based on **[P7]** and **[129–131, 134]**.

6.1 Existing Research on Wireless Self-backhauling in Ultra-Dense Cellular Networks

As discussed earlier in Chapter 2, wireless inband self-backhauling has recently been considered as a possible option for decreasing the cost of the densely deployed cellular networks of the future [42, 67, 98, 183, 217, 237, 268]. It would allow the ANs to backhaul all the data with a macro BS or a backhaul node (BN) without requiring any wired data link, significantly reducing the overall cost of the system. What is more, combining this concept with simultaneous UL and DL on the same frequency band, facilitated by the IBFD capability of the AN, would further improve the spectral efficiency of such a network. This makes IBFD self-backhauling an intriguing concept for the future 5G systems, as it facilitates higher data rates while also reducing the associated overall costs.

Such inband self-backhauling has been widely investigated in the earlier literature. Most works have considered a relay-type AN that is directly forwarding the signals transmitted by the UL UEs to the BN, or vice versa [41, 42, 83, 91, 97, 183, 217, 237, 268]. The reason for the popularity of this type of a scheme is likely the fact that such a

relay-type AN is more or less directly compatible with the existing networks, as it would essentially just extend the range of the macro BS or BN. This, on the other hand, will obviously result in increased data rates and better coverage.

In particular, in [83], the power control of such a relay-type AN is investigated, and the performance of both HD and IBFD operation modes is then compared. The obtained results indicate that the IBFD AN is capable of obtaining higher throughputs than the corresponding HD system, although a certain amount of SI suppression is obviously required. A similar analysis is performed in [42], where the BN is assumed to have a massive antenna array. There, the optimal power allocation for the BN and the AN is solved iteratively. The work in [268], on the other hand, investigates different beamforming solutions for a BN with massive antenna arrays, although no IBFD operation is assumed in any of the nodes therein.

Moreover, the effect of radio resource management (RRM) on the performance of the relay-type AN is investigated in [183]. There, the RRM tools are used to balance the SI with the other sources of interference, and the resulting solution is shown to outperform the HD benchmark scheme. In [97], the spectral efficiency of a similar system is maximized by solving the optimal power allocation for both IBFD and HD ANs. While the power allocation is solved in closed form for the HD case, only an algorithm is proposed for optimizing the transmit powers of the IBFD scenario. Also there the IBFD solution is shown to outperform the corresponding HD case.

The DL coverage of a relay-type self-backhauling AN is then analyzed in [217, 237]. The findings in [217] indicate that, while the throughput of the network with IBFD-capable ANs is almost doubled in comparison to the HD systems, the increased interference levels result in a somewhat smaller coverage. The results obtained in [237] suggest, on the other hand, that on a network level it may be better to have also some ANs that perform the self-backhauling on a different frequency band. This somewhat reduces the interference between the different backhaul links and the DL UEs. Finally, in [41], the throughput and outage probability of a relay-type IBFD AN is analyzed under an antenna selection scheme where individual TX and RX antennas are chosen at the AN based on a given criterion. Again, the IBFD AN is shown to usually outperform the corresponding HD AN, although this is not the case under all channel conditions.

All in all, even though different IBFD self-backhauling solutions for small cells have been widely investigated in the earlier literature, none of the above works have considered a scenario where also the UL and DL transmissions are performed simultaneously on the same center frequency. Considering the promising findings regarding the relay-type scenario where the DL and UL are separated either in time or in frequency, this means that the purely IBFD scheme analyzed in **[P7, 134]** and below is an intriguing option for further improving the spectral efficiency of these types of networks.

6.2 Analysis of Self-backhauling Full-Duplex Access Node with Massive Antenna Arrays

Article **[P7]** considers a system depicted in the top part of Fig. 6.1, where an IBFD-capable AN with massive TX and RX antenna arrays is serving HD-capable single-antenna UL and DL UEs while also using the same frequency band and antenna arrays

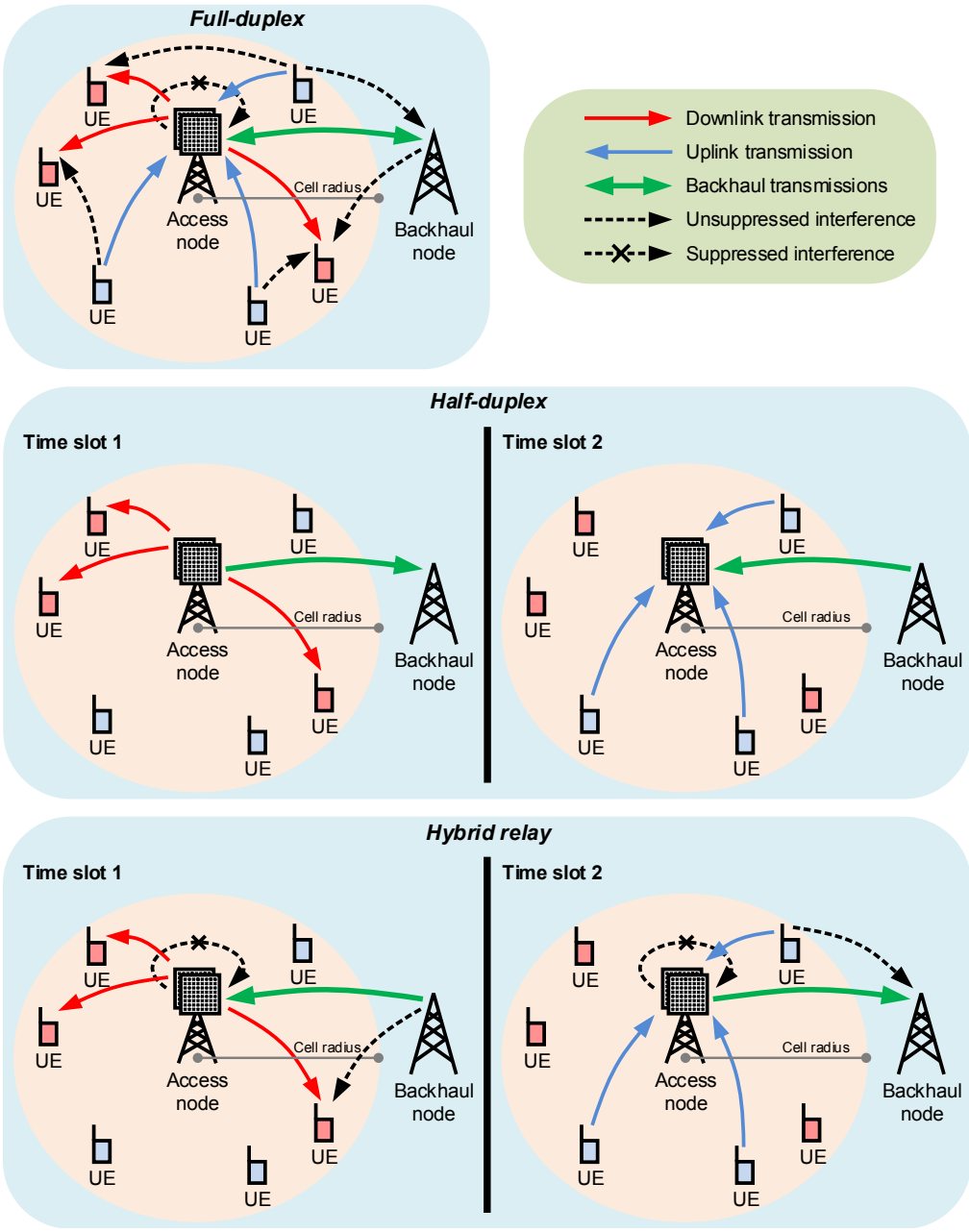


Figure 6.1: An illustration of the three considered communications schemes: the full-duplex scheme, the half-duplex scheme, and the hybrid relay scheme.

for backhauling the data. The wireless backhaul link is maintained by exchanging data with an IBFD-capable MIMO BN, which then further forwards the data either via a wired or wireless connection to the core network. The considered system is analyzed in terms of minimizing the transmit powers under some given QoS requirements, defined as minimum UL and DL data rates. Moreover, the IBFD solution is compared to two reference schemes, illustrated also in Fig. 6.1, which divide some or all of the communications tasks in time. To facilitate a straightforward comparison, all of the considered schemes are assumed to have the same amount of RF chains and TX/RX antennas at their disposal. Due to the various interference signals that are present in the schemes where the AN operates in IBFD mode, they will also have a fundamental boundary for the highest data rates that can be supported, which is also analyzed in detail in [**P7**] and in Section 6.2.3 below.

6.2.1 Different Communications Schemes

First, to establish the relationship between the transmit powers and the realized data rates, let us determine the SINRs of the different communicating parties. It is assumed that the AN has full channel state information (CSI) available and that it utilizes it to perform zero-forcing (ZF) beamforming, thereby separating or orthogonalizing the different spatial streams while also nulling the leakage to its own RX antennas to reduce the received SI power if necessary [176]. The assumption of full CSI is obviously an optimistic one for large antenna arrays, but it allows the derivation of the analytical data rate expressions in closed form, which, on the other hand, can be used to characterize the ultimate performance of the considered self-backhauling network.

Full-Duplex Scheme

The *full-duplex scheme*, depicted in the top part of Fig. 6.1, represents the scenario where all of the communications tasks are done simultaneously. While this in principle improves the spectral efficiency of the system, such a scheme suffers from various interference sources, most notably the SI and the UL-to-DL IUI. The latter is perhaps the more cumbersome issue as it cannot be canceled in a straightforward manner. Consequently, in this analysis, the IUI is mitigated only by choosing the DL and UL UEs from the opposite sides of the cell such that their mutual path loss is maximized [46, 60], although more advanced schemes have also been proposed in the literature [65, 114, 205].

For the full-duplex scheme, the SINR of the ith DL UE is shown in [**P7**] to be

$$sinr_i^{\mathrm{d,FD}} = \frac{\left(N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}} - N_{\mathrm{r}}\right) L_i^{\mathrm{d}} p_i^{\mathrm{d}}}{\sigma_{\mathrm{n}}^2 + L_i^{\mathrm{Bd}} P_{\mathrm{d}}^{\mathrm{B}} + \sum_{j=1}^{U} L_{ij}^{\mathrm{ud}} p_j^{\mathrm{u}}}, \tag{6.1}$$

where

- $N_{\mathrm{t}}/N_{\mathrm{r}}$ is the number of transmit/receive antennas in the AN;
- D/U is the number of DL/UL UEs;
- $M_{\mathrm{t}}^{\mathrm{B}}$ is the number of backhaul data streams transmitted by the AN;

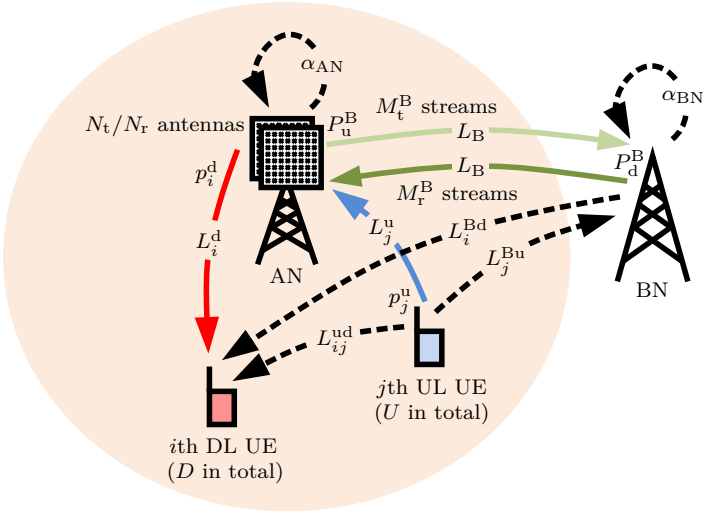


Figure 6.2: An illustration depicting the relevant symbols.

- p_i^{d} is the transmit power allocated for the ith DL signal;
- p_j^{u} is the transmit power of the jth UL UE;
- $P_{\mathrm{d}}^{\mathrm{B}}$ is the total transmit power of the BN;
- L_i^{d} is the path loss between the AN and the ith DL UE;
- L_i^{Bd} is the path loss between the BN and the ith DL UE;
- L_{ij}^{ud} is the path loss between the ith DL UE and the jth UL UE;
- σ_{n}^2 is the receiver noise floor.

The relevant symbols used throughout this chapter are also illustrated in Fig. 6.2, which provides a visual depiction of their meaning within the considered system. Investigating then (6.1), the loss of degrees-of-freedom due to directing nulls to the RX antennas is evident in the DL SINRs as the subtraction of the number of receive antennas in the numerator. Moreover, in addition to the UL-to-DL IUI, also the BN transmissions produce interference in the DL which somewhat reduces the SINRs.

Note that (6.1) and all the forthcoming SINR expressions are derived under the large-array assumption, i.e., the number of antennas in the AN being far greater than the number of transmitted or received signal streams [176, 258]. For the number of AN transmit/receive antennas, this is formally expressed as $N_{\mathrm{t}} \gg D+M_{\mathrm{t}}^{\mathrm{B}}$ and $N_{\mathrm{r}} \gg U+M_{\mathrm{r}}^{\mathrm{B}}$, where $M_{\mathrm{r}}^{\mathrm{B}}$ is the number of backhaul data streams received by the AN. For more details regarding the derivation of the SINRs, refer to **[P7]**.

Similarly, the SINR of the jth UL signal can consequently be written as **[P7]**

$$sinr_j^{\mathrm{u,FD}} = \frac{\left(N_{\mathrm{r}} - U - M_{\mathrm{r}}^{\mathrm{B}}\right) L_j^{\mathrm{u}} p_j^{\mathrm{u}}}{\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{AN}} \left(P_{\mathrm{u}}^{\mathrm{B}} + \sum_{i=1}^{D} p_i^{\mathrm{d}}\right)}, \tag{6.2}$$

where

- L_j^{u} is the path loss between the AN and the jth UL UE;
- α_{AN} is the total amount of SI suppression in the AN, consisting of the ZF nulling and potentially other active SI cancellation stages;
- $P_{\mathrm{u}}^{\mathrm{B}}$ is the total transmit power allocated for backhauling the UL data.

Moreover, these symbols are also illustrated in Fig. 6.2. Now, since the SI nulling is done in the transmitter, no degrees-of-freedom are lost due to that when receiving data, as can be observed in (6.2). However, some residual SI still remains, which decreases the SINRs of the UL signals.

Then, to quantify the backhauling capability of the AN, the SINR expressions for the backhaul link are also required. Under the assumptions made in this analysis, the SINRs of the different backhaul data streams are in fact identical, the individual SINRs for DL and UL backhaul links being as follows [**P7**]:

$$sinr_{\mathrm{B}}^{\mathrm{d,FD}} = \frac{\left(N_{\mathrm{r}} - U - M_{\mathrm{r}}^{\mathrm{B}}\right) L_{\mathrm{B}} P_{\mathrm{d}}^{\mathrm{B}} / M_{\mathrm{r}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{AN}} \left(P_{\mathrm{u}}^{\mathrm{B}} + \sum_{i=1}^{D} p_i^{\mathrm{d}}\right)}, \tag{6.3}$$

$$sinr_{\mathrm{B}}^{\mathrm{u,FD}} = \frac{\left(N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}} - N_{\mathrm{r}}\right) L_{\mathrm{B}} P_{\mathrm{u}}^{\mathrm{B}} / M_{\mathrm{t}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{BN}} P_{\mathrm{d}}^{\mathrm{B}} + \sum_{j=1}^{U} L_j^{\mathrm{Bu}} p_j^{\mathrm{u}}}, \tag{6.4}$$

where

- L_{B} is the path loss between the AN and the BN;
- α_{BN} is the total amount of SI suppression in the BN;
- L_j^{Bu} is the path loss between the BN and the jth UL UE.

Here, (6.3) represents the SINR of the backhaul signals received by the AN, while (6.4) is the SINR of the signals received by the BN. Both of these are decreased by the residual SI in the respective receivers, while the latter is also reduced by the interference caused by the simultaneous UL transmissions.

Assuming then Gaussian distributed noise and interference within the system, the corresponding sum-rates or sum-capacities for the different links can simply be expressed as follows [**P7**]:

$$R^{\mathrm{d}} = \sum_{i=1}^{D} R_i^{\mathrm{d}}, \quad \text{where} \quad R_i^{\mathrm{d}} = \log_2\left(1 + sinr_i^{\mathrm{d,FD}}\right), \tag{6.5}$$

$$R^{\mathrm{u}} = \sum_{j=1}^{U} R_j^{\mathrm{u}}, \quad \text{where} \quad R_j^{\mathrm{u}} = \log_2\left(1 + sinr_j^{\mathrm{u,FD}}\right), \tag{6.6}$$

$$R_{\mathrm{B}}^{\mathrm{d}} = M_{\mathrm{r}}^{\mathrm{B}} \log_2\left(1 + sinr_{\mathrm{B}}^{\mathrm{d,FD}}\right), \tag{6.7}$$

$$R_{\mathrm{B}}^{\mathrm{u}} = M_{\mathrm{t}}^{\mathrm{B}} \log_2\left(1 + sinr_{\mathrm{B}}^{\mathrm{u,FD}}\right). \tag{6.8}$$

The expressions in (6.5)–(6.8) can then be used to obtain the closed-form solution for the optimal transmit powers under the given QoS conditions, as shown in Section 6.2.2. In particular, the DL and UL data rates in (6.5) and (6.6) are constrained by the QoS requirements, while the backhauling data rates in (6.7) and (6.8) are constrained by the realized DL and UL data rates as the AN must be capable of backhauling all the data.

Half-Duplex Scheme

Considering then the *half-duplex scheme* depicted in the middle part of Fig. 6.1, now transmission and reception within an individual device are separated in time, i.e., the system essentially operates in TDD mode. The relative lengths of the resulting two time slots are controlled by the duplexing parameter η, which determines the proportion of time spent in the DL mode, the rest of the time being then allocated for the UL mode. Moreover, to avoid simultaneous transmission and reception, the AN backhauls the UL data during the DL time slot, and correspondingly the DL data during the UL time slot. Although this introduces some delay compared to the full-duplex scheme, it can be expected to be negligible with a sufficiently short duration for the individual time slot.

Now, the SINRs of the different links can be expressed as follows:

$$sinr_i^{\mathrm{d,HD}} = \frac{\left(N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}}\right) L_i^{\mathrm{d}} p_i^{\mathrm{d}}}{\sigma_{\mathrm{n}}^2}, \tag{6.9}$$

$$sinr_j^{\mathrm{u,HD}} = \frac{\left(N_{\mathrm{r}} - U - M_{\mathrm{r}}^{\mathrm{B}}\right) L_j^{\mathrm{u}} p_j^{\mathrm{u}}}{\sigma_{\mathrm{n}}^2}, \tag{6.10}$$

$$sinr_{\mathrm{B}}^{\mathrm{d,HD}} = \frac{\left(N_{\mathrm{r}} - U - M_{\mathrm{r}}^{\mathrm{B}}\right) L_{\mathrm{B}} P_{\mathrm{d}}^{\mathrm{B}} / M_{\mathrm{r}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2}, \tag{6.11}$$

$$sinr_{\mathrm{B}}^{\mathrm{u,HD}} = \frac{\left(N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}}\right) L_{\mathrm{B}} P_{\mathrm{u}}^{\mathrm{B}} / M_{\mathrm{t}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2}. \tag{6.12}$$

Note that in this scheme no degrees-of-freedom are lost due to the nulling of the RX antennas as no SI is produced. Moreover, due to the TDD operation, the system does not suffer from any interference sources under the adopted modeling principles. The resulting data rates are then

$$R^{\mathrm{d}} = \sum_{i=1}^{D} R_i^{\mathrm{d}}, \quad \text{where} \quad R_i^{\mathrm{d}} = \eta \log_2\left(1 + sinr_i^{\mathrm{d,HD}}\right), \tag{6.13}$$

$$R^{\mathrm{u}} = \sum_{j=1}^{U} R_j^{\mathrm{u}}, \quad \text{where} \quad R_j^{\mathrm{u}} = (1-\eta) \log_2\left(1 + sinr_j^{\mathrm{u,HD}}\right), \tag{6.14}$$

$$R_{\mathrm{B}}^{\mathrm{d}} = (1-\eta)\, M_{\mathrm{r}}^{\mathrm{B}} \log_2\left(1 + sinr_{\mathrm{B}}^{\mathrm{d,HD}}\right), \tag{6.15}$$

$$R_{\mathrm{B}}^{\mathrm{u}} = \eta M_{\mathrm{t}}^{\mathrm{B}} \log_2\left(1 + sinr_{\mathrm{B}}^{\mathrm{u,HD}}\right). \tag{6.16}$$

Hence, while there are no interference sources in this scheme, the data rates are decreased by the need to divide the time between transmission and reception in the AN.

Hybrid Relay Scheme

The *hybrid relay scheme*, depicted in the bottom part of Fig. 6.1, is essentially a trade-off between TDD and IBFD operation. Namely, in this scheme, the AN acts as a relay between the BN and the UEs, meaning that during the first time slot it forwards the DL data from the BN to the UEs, after which it forwards the UL data from the UEs to the BN. Again, the duplexing parameter η specifies the proportion of time spent in the DL time slot. In this scheme, the AN takes some advantage of its IBFD capability while still avoiding the UL-to-DL IUI, which is a significant issue in the full-duplex scheme.

For the hybrid relay scheme, the different SINRs are

$$sinr_i^{\mathrm{d,RL}} = \frac{(N_\mathrm{t} - D - N_\mathrm{r})\, L_i^\mathrm{d} p_i^\mathrm{d}}{\sigma_\mathrm{n}^2 + L_i^\mathrm{Bd} P_\mathrm{d}^\mathrm{B}}, \tag{6.17}$$

$$sinr_j^{\mathrm{u,RL}} = \frac{(N_\mathrm{r} - U)\, L_j^\mathrm{u} p_j^\mathrm{u}}{\sigma_\mathrm{n}^2 + \alpha_\mathrm{AN} P_\mathrm{u}^\mathrm{B}}, \tag{6.18}$$

$$sinr_\mathrm{B}^{\mathrm{d,RL}} = \frac{\left(N_\mathrm{r} - M_\mathrm{r}^\mathrm{B}\right) L_\mathrm{B} P_\mathrm{d}^\mathrm{B} / M_\mathrm{r}^\mathrm{B}}{\sigma_\mathrm{n}^2 + \alpha_\mathrm{AN} \sum_{i=1}^{D} p_i^\mathrm{d}}, \tag{6.19}$$

$$sinr_\mathrm{B}^{\mathrm{u,RL}} = \frac{\left(N_\mathrm{t} - M_\mathrm{t}^\mathrm{B} - N_\mathrm{r}\right) L_\mathrm{B} P_\mathrm{u}^\mathrm{B} / M_\mathrm{t}^\mathrm{B}}{\sigma_\mathrm{n}^2 + \sum_{j=1}^{U} L_j^\mathrm{Bu} p_j^\mathrm{u}}. \tag{6.20}$$

Correspondingly, the data rates are now as follows:

$$R^\mathrm{d} = \sum_{i=1}^{D} R_i^\mathrm{d}, \quad \text{where} \quad R_i^\mathrm{d} = \eta \log_2\left(1 + sinr_i^{\mathrm{d,RL}}\right), \tag{6.21}$$

$$R^\mathrm{u} = \sum_{j=1}^{U} R_j^\mathrm{u}, \quad \text{where} \quad R_j^\mathrm{u} = (1-\eta) \log_2\left(1 + sinr_j^{\mathrm{u,RL}}\right), \tag{6.22}$$

$$R_\mathrm{B}^\mathrm{d} = \eta M_\mathrm{r}^\mathrm{B} \log_2\left(1 + sinr_\mathrm{B}^{\mathrm{d,RL}}\right), \tag{6.23}$$

$$R_\mathrm{B}^\mathrm{u} = (1-\eta)\, M_\mathrm{t}^\mathrm{B} \log_2\left(1 + sinr_\mathrm{B}^{\mathrm{u,RL}}\right). \tag{6.24}$$

As can be observed, the hybrid relay scheme loses some degrees-of-freedom due to having to direct nulls to the RX antennas in the AN while also suffering from the residual SI and the interference between the BN and the UEs. However, it avoids the cumbersome UL-to-DL IUI entirely, and hence this scheme likely represents a sensible trade-off between maximizing the temporal efficiency and minimizing the amount of interference.

6.2.2 Transmit Power Optimization

As already mentioned, the objective of the following analysis is to minimize the transmit powers of the different communicating parties while still ensuring that the minimum QoS is fulfilled. In particular, it is assumed that the optimal transmit powers are calculated by the AN, who then forwards them to the relevant parties. Taking into account the

QoS and the backhauling requirements, this optimization problem can be formulated for all the communications schemes as follows:

$$\begin{aligned} \underset{\mathbf{p},\, P_{\mathrm{d}}^{\mathrm{B}},\, P_{\mathrm{u}}^{\mathrm{B}}}{\text{minimize}} \quad & \left(\mathbf{1}_{D+U}^T \mathbf{p} + P_{\mathrm{d}}^{\mathrm{B}} + P_{\mathrm{u}}^{\mathrm{B}}\right) \\ \text{subject to} \quad & \text{C1: } R_i^{\mathrm{d}} \geq \rho_{\mathrm{d}},\ i = 1, \ldots, D, \\ & \text{C2: } R_j^{\mathrm{u}} \geq \rho_{\mathrm{u}},\ j = 1, \ldots, U, \\ & \text{C3: } R_{\mathrm{B}}^{\mathrm{d}} \geq \sum_{i=1}^{D} R_i^{\mathrm{d}}, \\ & \text{C4: } R_{\mathrm{B}}^{\mathrm{u}} \geq \sum_{j=1}^{U} R_j^{\mathrm{u}}, \end{aligned} \tag{6.25}$$

where $\mathbf{p} = \begin{bmatrix} \mathbf{p}_{\mathrm{d}}^T & \mathbf{p}_{\mathrm{u}}^T \end{bmatrix}^T$, $\mathbf{p}_{\mathrm{d}}$ and $\mathbf{p}_{\mathrm{u}}$ are column vectors defined as $\{\mathbf{p}_{\mathrm{d}}\}_i = p_i^{\mathrm{d}}$ and $\{\mathbf{p}_{\mathrm{u}}\}_j = p_j^{\mathrm{u}}$, respectively, while ρ_{d} and ρ_{u} are the corresponding per-UE minimum data rate requirements

Here, the constraints C1 and C2 guarantee that the given QoS is fulfilled, while the constraints C3 and C4 ensure the backhauling capability of the AN. Note that the optimization problem does not explicitly require the optimal transmit powers to be positive. This is intentional since obtaining negative optimal transmit powers indicates that the given QoS requirements cannot be fulfilled with any finite transmit powers and the system is consequently infeasible, which is a meaningful finding in itself. These infeasible system scenarios are characterized in detail in Section 6.2.3 below.

To facilitate a more straightforward notation in the continuation, the different path losses are grouped into vectors and matrices as follows:

$$\begin{aligned} \{\mathbf{q}_{\mathrm{d}}\}_i &= 1/L_i^{\mathrm{d}}, \\ \{\mathbf{q}_{\mathrm{u}}\}_j &= 1/L_j^{\mathrm{u}}, \\ \{\mathbf{q}_{\mathrm{B/d}}\}_i &= L_i^{\mathrm{Bd}}/L_i^{\mathrm{d}}, \\ \{\mathbf{L}_{\mathrm{Bu}}\}_j &= L_j^{\mathrm{Bu}}, \\ \{\mathbf{L}_{\mathrm{ud}}\}_{ij} &= L_{ij}^{\mathrm{ud}}, \\ \mathbf{L}_{\mathrm{ud}}^{\mathrm{d}} &= \mathbf{L}_{\mathrm{ud}} \circ \mathbf{q}_{\mathrm{d}} \mathbf{1}_U^T. \end{aligned}$$

In addition, the sums of some of the above path loss vectors and matrices are defined as

$$\begin{aligned} S_{\mathrm{d}} &= \mathbf{1}_D^T \mathbf{q}_{\mathrm{d}}, \\ S_{\mathrm{B/d}} &= \mathbf{1}_D^T \mathbf{q}_{\mathrm{B/d}}, \\ S_{\mathrm{B/u}} &= \mathbf{L}_{\mathrm{Bu}}^T \mathbf{q}_{\mathrm{u}}, \\ S_{\mathrm{ud}} &= \mathbf{1}_D^T \mathbf{L}_{\mathrm{ud}}^{\mathrm{d}} \mathbf{q}_{\mathrm{u}}. \end{aligned}$$

With the help of these variables, the solutions to the above optimization problem are next given in closed form for each communications scheme.

Full-Duplex Scheme

The optimal DL and UL transmit powers for the full-duplex scheme can be written as follows [**P7**]:

$$\mathbf{p}^{\mathrm{opt}} = \begin{bmatrix} \mathbf{p}_{\mathrm{d}}^{\mathrm{opt}} \\ \mathbf{p}_{\mathrm{u}}^{\mathrm{opt}} \end{bmatrix} = \begin{bmatrix} \frac{\sigma_{\mathrm{n}}^2 \gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}} \mathbf{q}_{\mathrm{d}} + \frac{\sigma_{\mathrm{n}}^2 \left(\frac{1+S_{\mathrm{d}}\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}} + \gamma_{\mathrm{t}}^{\mathrm{B}} \right)}{\theta^{\mathrm{FD}}(1-\gamma^{\mathrm{B}})} \left(\gamma_{\mathrm{d}} \gamma_{\mathrm{r}}^{\mathrm{B}} \mathbf{q}_{\mathrm{B/d}} + \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}}}{\alpha_{\mathrm{AN}}} \mathbf{L}_{\mathrm{ud}}^{\mathrm{d}} \mathbf{q}_{\mathrm{u}} \right) \\ \frac{\sigma_{\mathrm{n}}^2 \gamma_{\mathrm{u}} \left(\frac{1+S_{\mathrm{d}}\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}} + \gamma_{\mathrm{t}}^{\mathrm{B}} \right)}{\theta^{\mathrm{FD}}(1-\gamma^{\mathrm{B}})} \mathbf{q}_{\mathrm{u}} \end{bmatrix}, \tag{6.26}$$

$$\left(P_{\mathrm{d}}^{\mathrm{B}}\right)^{\mathrm{opt}} = \frac{\alpha_{\mathrm{AN}}\gamma_{\mathrm{r}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}} \mathbf{1}_D^T \mathbf{p}_{\mathrm{d}}^{\mathrm{opt}} + \frac{\alpha_{\mathrm{AN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}} \mathbf{L}_{\mathrm{Bu}}^T \mathbf{p}_{\mathrm{u}}^{\mathrm{opt}} + \frac{\left(1+\alpha_{\mathrm{AN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\right)\gamma_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{1-\gamma^{\mathrm{B}}}, \tag{6.27}$$

$$\left(P_{\mathrm{u}}^{\mathrm{B}}\right)^{\mathrm{opt}} = \frac{\gamma^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}} \mathbf{1}_D^T \mathbf{p}_{\mathrm{d}}^{\mathrm{opt}} + \frac{\gamma_{\mathrm{t}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}} \mathbf{L}_{\mathrm{Bu}}^T \mathbf{p}_{\mathrm{u}}^{\mathrm{opt}} + \frac{\left(1+\alpha_{\mathrm{BN}}\gamma_{\mathrm{r}}^{\mathrm{B}}\right)\gamma_{\mathrm{t}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{1-\gamma^{\mathrm{B}}}, \tag{6.28}$$

when each transmit power is positive; otherwise the QoS requirements cannot be fulfilled and the system is infeasible. Here, the following auxiliary variables have been used:

$$\gamma_{\mathrm{d}} = \frac{\left(2^{\rho_{\mathrm{d}}} - 1\right)\alpha_{\mathrm{AN}}}{N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}} - N_{\mathrm{r}}}, \tag{6.29}$$

$$\gamma_{\mathrm{u}} = \frac{\left(2^{\rho_{\mathrm{u}}} - 1\right)\alpha_{\mathrm{AN}}}{N_{\mathrm{r}} - U - M_{\mathrm{r}}^{\mathrm{B}}}, \tag{6.30}$$

$$\gamma_{\mathrm{r}}^{\mathrm{B}} = \frac{\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}} - 1\right) M_{\mathrm{r}}^{\mathrm{B}}}{L_{\mathrm{B}}\left(N_{\mathrm{r}} - U - M_{\mathrm{r}}^{\mathrm{B}}\right)}, \tag{6.31}$$

$$\gamma_{\mathrm{t}}^{\mathrm{B}} = \frac{\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}} - 1\right) M_{\mathrm{t}}^{\mathrm{B}}}{L_{\mathrm{B}}\left(N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}} - N_{\mathrm{r}}\right)}, \tag{6.32}$$

$$\gamma^{\mathrm{B}} = \alpha_{\mathrm{AN}}\alpha_{\mathrm{BN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}, \tag{6.33}$$

$$\theta^{\mathrm{FD}} = 1 - \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}} S_{\mathrm{B/d}}}{1-\gamma^{\mathrm{B}}} - \frac{\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}} S_{\mathrm{B/u}}}{1-\gamma^{\mathrm{B}}} - \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}} S_{\mathrm{ud}}}{\alpha_{\mathrm{AN}}\left(1-\gamma^{\mathrm{B}}\right)}, \tag{6.34}$$

where especially the variables γ^{B} and θ^{FD} are closely related to the fundamental feasibility of the full-duplex scheme, as will be shown in Section 6.2.3. Although the detailed proof of the above solution is omitted as it is already available in [**P7**], it is obtained in principle by presenting the optimization problem in (6.25) as a system of linear equations. Then, solving the optimal transmit powers is essentially reduced to a matrix inversion, which in this case can be done in closed form. For further details regarding the derivation of (6.26)–(6.28), refer to [**P7**].

Half-Duplex Scheme

The optimal transmit powers for the half-duplex scheme are correspondingly [**P7**]

$$\mathbf{p}^{\mathrm{opt}} = \begin{bmatrix} \mathbf{p}_{\mathrm{d}}^{\mathrm{opt}} \\ \mathbf{p}_{\mathrm{u}}^{\mathrm{opt}} \end{bmatrix} = \begin{bmatrix} \frac{\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}}-1\right)\sigma_n^2}{N_{\mathrm{t}}-D-M_{\mathrm{t}}^{\mathrm{B}}}\mathbf{q}_{\mathrm{d}} \\ \frac{\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}}-1\right)\sigma_{\mathrm{n}}^2}{N_{\mathrm{r}}-U-M_{\mathrm{r}}^{\mathrm{B}}}\mathbf{q}_{\mathrm{u}} \end{bmatrix}, \tag{6.35}$$

$$\left(P_{\mathrm{d}}^{\mathrm{B}}\right)^{\mathrm{opt}} = \frac{\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}(1-\eta)}}-1\right) M_{\mathrm{r}}^{\mathrm{B}}\sigma_n^2}{L_{\mathrm{B}}\left(N_{\mathrm{r}}-U-M_{\mathrm{r}}^{\mathrm{B}}\right)}, \tag{6.36}$$

$$\left(P_{\mathrm{u}}^{\mathrm{B}}\right)^{\mathrm{opt}} = \frac{\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}\eta}}-1\right) M_{\mathrm{t}}^{\mathrm{B}}\sigma_n^2}{L_{\mathrm{B}}\left(N_{\mathrm{t}}-D-M_{\mathrm{t}}^{\mathrm{B}}\right)}. \tag{6.37}$$

In this case, these optimal solutions can be obtained by simply choosing the transmit powers such that the constraints C1–C4 are fulfilled with equality. As the transmit powers do not affect the interference levels, this is then directly the solution to the problem. For this same reason, the half-duplex scheme can always fulfill the QoS requirements and is consequently feasible under all circumstances.

However, it should be noted that the solution specified by (6.35)–(6.37) is only for a given duplexing parameter. Hence, unless the lengths of the duplexing periods are fixed by the system, the optimization must also be performed with respect to η. In other words, the sum of the optimal transmit powers in (6.35)–(6.37) must be minimized also with η as the argument. In [**P7**], it is shown that this sum transmit power is in fact a convex function of η, meaning that the optimal value can be found at the zero-point of its derivative. This problem has no closed-form solution, and hence it must be solved numerically when evaluating the optimal transmit powers.

Hybrid Relay Scheme

Finally, the optimal transmit powers for the hybrid relay scheme can be expressed as follows:

$$\mathbf{p}^{\mathrm{opt}} = \begin{bmatrix} \mathbf{p}_{\mathrm{d}}^{\mathrm{opt}} \\ \mathbf{p}_{\mathrm{u}}^{\mathrm{opt}} \end{bmatrix} = \begin{bmatrix} \frac{\gamma_{\mathrm{d}}\sigma_{\mathrm{n}}^2}{\alpha_{\mathrm{AN}}}\left(\mathbf{q}_{\mathrm{d}} + \frac{\gamma_{\mathrm{r}}^{\mathrm{B}}(1+S_{\mathrm{d}}\gamma_{\mathrm{d}})}{1-\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}\mathbf{q}_{\mathrm{B/d}}\right) \\ \frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{u}}}{\alpha_{\mathrm{AN}}}\left(\frac{1+\alpha_{\mathrm{AN}}\gamma_{\mathrm{t}}^{\mathrm{B}}}{1-\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\mathbf{q}_{\mathrm{u}}\right) \end{bmatrix}, \tag{6.38}$$

$$\left(P_{\mathrm{d}}^{\mathrm{B}}\right)^{\mathrm{opt}} = \gamma_{\mathrm{r}}^{\mathrm{B}}\left(\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{AN}}\mathbf{1}_D^T\mathbf{p}_{\mathrm{d}}^{\mathrm{opt}}\right), \tag{6.39}$$

$$\left(P_{\mathrm{u}}^{\mathrm{B}}\right)^{\mathrm{opt}} = \gamma_{\mathrm{t}}^{\mathrm{B}}\left(\sigma_{\mathrm{n}}^2 + \mathbf{L}_{\mathrm{Bu}}^T\mathbf{p}_{\mathrm{u}}^{\mathrm{opt}}\right), \tag{6.40}$$

when the individual transmit powers are positive; otherwise the system is infeasible. The auxiliary variables are now defined as

$$\gamma_{\mathrm{d}} = \frac{\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}} - 1\right)\alpha_{\mathrm{AN}}}{N_{\mathrm{t}} - D - N_{\mathrm{r}}}, \tag{6.41}$$

$$\gamma_{\mathrm{u}} = \frac{\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}} - 1\right)\alpha_{\mathrm{AN}}}{N_{\mathrm{r}} - U}, \tag{6.42}$$

$$\gamma_{\mathrm{r}}^{\mathrm{B}} = \frac{\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}\eta}} - 1\right)M_{\mathrm{r}}^{\mathrm{B}}}{L_{\mathrm{B}}\left(N_{\mathrm{r}} - M_{\mathrm{r}}^{\mathrm{B}}\right)}, \tag{6.43}$$

$$\gamma_{\mathrm{t}}^{\mathrm{B}} = \frac{\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}(1-\eta)}} - 1\right)M_{\mathrm{t}}^{\mathrm{B}}}{L_{\mathrm{B}}\left(N_{\mathrm{t}} - M_{\mathrm{t}}^{\mathrm{B}} - N_{\mathrm{r}}\right)}. \tag{6.44}$$

This optimal solution is obtained in a similar manner as that of the full-duplex scheme, i.e., by solving a system of linear equations, which directly follows from the original optimization problem. For further details, refer to [**P7**].

Similar to the half-duplex scheme, the above solution must still be optimized with respect to the duplexing parameter. Again, this can be done by using the sum of the optimal transmit powers in (6.38)–(6.40) as the objective function, and determining the value of η that minimizes it [**P7**]. Furthermore, as there is no closed-form solution for the optimal η, it is numerically optimized in the forthcoming numerical results.

Computational Complexity of the Optimization Procedure

Let us then briefly consider the computational requirements of calculating the optimal transmit powers within the AN, using the solutions defined above. Thanks to the closed-form expressions, no iterations are required in determining the transmit power allocation, which results in a rather straightforward optimization framework in terms of the computations. What is more, the closed-form solutions involve only additions, multiplications, and divisions, further reducing the computational cost. Assuming that the system is operating under static conditions where only the path losses are changing, many of the terms can also be precomputed. This is a reasonable assumption as it can be expected that the frequency with which the system parameters are changed, if changed at all, is much lower than the frequency of estimating the channels and path losses. Moreover, in this simple analysis, the additions are omitted since the multiplications and divisions dominate the computational cost [92], while the number of divisions is also kept as small as possible. It should also be emphasized that the power allocation scheme proposed in this thesis has been derived under idealized assumptions to provide information about the ultimate performance of the system, and hence it should not be considered a practical algorithm as such. Nevertheless, the complexities of the different communications schemes can still be compared to gain knowledge regarding their relative computational requirements.

Table 6.1: The computational requirements of transmit power optimization for the different communications schemes.

	Number of multiplications	**Number of divisions**
Full-duplex scheme	$5D + 3U + 2DU + 27$	$D + U + 3$
Half-duplex scheme	2	$D + U + 1$
Hybrid relay scheme	$3D + 3U + 12$	$D + U + 3$

Starting from the full-duplex scheme, the auxiliary variables γ_{d} and γ_{u} can be precomputed, alongside with all the other terms consisting of only these and other static system parameters. The vectors, matrices, and variables containing the path losses, on the other hand, must obviously be computed in real-time in order to obtain the optimal transmit powers. It can easily be deduced that computing the elements of the different path loss vectors and matrices in $\mathbf{q}_{\mathrm{d}}$, $\mathbf{q}_{\mathrm{u}}$, $\mathbf{q}_{\mathrm{B/d}}$, and $\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}$, together with the quantities in $\gamma_{\mathrm{r}}^{\mathrm{B}}$, $\gamma_{\mathrm{t}}^{\mathrm{B}}$, γ^{B}, S_{d}, $S_{\mathrm{B/d}}$, $S_{\mathrm{B/u}}$, and S_{ud}, requires $D + U + 2DU + 4$ multiplications and $D + U + 1$ divisions in total (the number of multiplications is increased by calculating also the vector $\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}\mathbf{q}_{\mathrm{u}}$ for later use). Having computed the quantities of these variables, the value of the variable θ^{FD} is obtained with eight multiplications and a single division, recalling again that the terms containing only the static system parameters have been precomputed. Then, the optimal DL transmit powers can be computed with $4D + 2$ multiplications and a single division, while the corresponding optimal UL transmit powers require $U + 1$ multiplications. Finally, utilizing the optimal DL and UL transmit powers, the backhaul-related transmit powers can be calculated with $U + 12$ multiplications. Therefore, in total, the full-duplex scheme requires $5D + 3U + 2DU + 27$ multiplications and $D + U + 3$ divisions for calculating the optimal transmit powers.

Considering then the half-duplex scheme, now the optimal DL and UL transmit powers are obtained simply with D and U divisions, respectively, while solving the backhaul-related transmit powers requires two multiplications and a single division. Hence, the total number of computations required for solving the optimal transmit powers is two multiplications and $D + U + 1$ divisions. However, it should be noted that this is only valid for a given duplexing parameter η, whose optimization will also require some computations. However, as it likely suffices to optimize the duplexing parameter much less frequently than the transmit powers, the computational cost of finding the optimal η is not considered in this thesis.

In the hybrid relay scheme, it takes $D + U + 2$ multiplications and $D + U + 1$ divisions to compute the path loss dependent vectors, matrices and variables in $\mathbf{q}_{\mathrm{d}}$, $\mathbf{q}_{\mathrm{u}}$, $\mathbf{q}_{\mathrm{B/d}}$, $\gamma_{\mathrm{r}}^{\mathrm{B}}$, $\gamma_{\mathrm{t}}^{\mathrm{B}}$, S_{d}, $S_{\mathrm{B/d}}$, and $S_{\mathrm{B/u}}$. Having obtained these quantities, the optimal DL and UL transmit powers can be computed with $2D + U + 7$ multiplications and two divisions while the backhaul-related transmit powers require $U + 3$ multiplications. As a result, the total number of computations needed for solving the optimal transmit power allocation is $3D + 3U + 12$ multiplications and $D + U + 3$ divisions. Similar to the half-duplex scheme, these computations do not include the cost of finding the optimal η. Nevertheless, as already mentioned, it is unlikely that the duplexing parameter has to be optimized as frequently as the transmit powers, reducing its contribution to the overall computational complexity.

The computational costs of transmit power optimization for all the schemes are collected in Table 6.1. As can be observed, the full-duplex scheme requires the most computations, while the half-duplex scheme is computationally the simplest option. However, as will be observed in Section 6.2.4, the benefit of the higher complexity of the full-duplex scheme is the lower overall transmit power consumption. It should also be noted that these results are only valid when the transmit powers are calculated sequentially, i.e., when the quantities calculated earlier are used also in the later computation stages. More multiplications would be required if the optimal transmit powers were computed in parallel, although then the latency of the optimization procedure would correspondingly be decreased.

6.2.3 Feasibility Analysis

As already discussed, the feasibility of the system is determined by the sign of the optimal transmit powers; if any of the optimal transmit powers are negative, the given QoS requirements cannot be fulfilled and the system is infeasible. On the other hand, if all of the optimal transmit powers are positive and finite, the system is feasible. This stems simply from the physical interpretation of a transmit power, which cannot be negative or infinite.

In general, the possibility of the system being infeasible is caused by the interference sources. Namely, as a result of the different interference links, using higher transmit powers at certain nodes increases the useful signal power of some communicating parties while also increasing the interference level of some other parties. This naturally results in an upper bound for the obtainable data rates both in the DL and in the UL as some of the SINRs saturate when increasing the transmit powers sufficiently high [**P7, 130**]. Consequently, if the given data rate requirements surpass this fundamental upper bound, the system is infeasible as the required data rates cannot be reached with any finite transmit powers. Both the full-duplex scheme and the hybrid relay scheme suffer from this potential infeasibility as they contain various interference links. However, the half-duplex scheme is always feasible as it does not produce any internal interference under the considered assumptions.

Below, the closed-form boundary conditions for the feasibility of the full-duplex and hybrid relay schemes are defined and discussed. These boundaries correspond essentially to a case where the transmit powers tend to infinity, and consequently they are very fundamental in nature. The corresponding boundaries under limited transmit powers would obviously be somewhat stricter.

Feasibility of the Full-Duplex Scheme

It can be shown that the full-duplex scheme is feasible under the following conditions:

$$\begin{cases} \dfrac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}} S_{\mathrm{B/d}}}{1-\gamma^{\mathrm{B}}} + \dfrac{\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}} S_{\mathrm{B/u}}}{1-\gamma^{\mathrm{B}}} + \dfrac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}} S_{\mathrm{ud}}}{\alpha_{\mathrm{AN}}\left(1-\gamma^{\mathrm{B}}\right)} < 1, \\ \gamma^{\mathrm{B}} < 1. \end{cases} \tag{6.45}$$

Note that the first condition is simply $\theta^{\mathrm{FD}} > 0$ rewritten in a different form. The proof of these feasibility conditions can be found in [**P7**].

The exact condition in (6.45) can be simplified by noting that the term $\alpha_{\mathrm{AN}}\alpha_{\mathrm{BN}}$ is extremely small, meaning that $\gamma^{\mathrm{B}} \approx 0$. Hence, the condition $\gamma^{\mathrm{B}} < 1$ is practically always fulfilled, while the other condition reduces to

$$\alpha_{\mathrm{AN}} \left(\frac{(2^{\rho_{\mathrm{u}}}-1)(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}}-1)M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}} + (2^{\rho_{\mathrm{d}}}-1)(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}}-1)M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{L_{\mathrm{B}}\left(N_{\mathrm{t}}-D-M_{\mathrm{t}}^{\mathrm{B}}-N_{\mathrm{r}}\right)\left(N_{\mathrm{r}}-U-M_{\mathrm{r}}^{\mathrm{B}}\right)} + \frac{(2^{\rho_{\mathrm{u}}}-1)(2^{\rho_{\mathrm{d}}}-1)S_{\mathrm{ud}}}{\left(N_{\mathrm{t}}-D-M_{\mathrm{t}}^{\mathrm{B}}-N_{\mathrm{r}}\right)\left(N_{\mathrm{r}}-U-M_{\mathrm{r}}^{\mathrm{B}}\right)} \right) < 1. \tag{6.46}$$

From this simplified expression, the maximum achievable DL data rate requirement can be obtained by solving its root with respect to ρ_{d}. This can only be done numerically, and consequently no closed-form solution for the achievable DL data rate region exists. As opposed to this, the SI cancellation requirement under some given system parameters and QoS requirements can easily be obtained from (6.46) in closed form as follows:

$$\begin{aligned} \alpha_{\mathrm{AN}}^{dB} < {} & 10\log_{10}\left[L_{\mathrm{B}}\left(N_{\mathrm{t}}-D-M_{\mathrm{t}}^{\mathrm{B}}-N_{\mathrm{r}}\right)\left(N_{\mathrm{r}}-U-M_{\mathrm{r}}^{\mathrm{B}}\right)\right] \\ & - 10\log_{10}\Bigg[(2^{\rho_{\mathrm{u}}}-1)\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}}-1\right)M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}} \\ & \quad + (2^{\rho_{\mathrm{d}}}-1)\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}}-1\right)M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}} \\ & \quad + (2^{\rho_{\mathrm{u}}}-1)(2^{\rho_{\mathrm{d}}}-1)L_{\mathrm{B}}S_{\mathrm{ud}}\Bigg]. \end{aligned} \tag{6.47}$$

Note that this represents the minimum SI cancellation requirement to make the full-duplex scheme feasible when the transmit powers tend to infinity. Therefore, more SI cancellation is obviously needed when the transmit powers are limited.

Feasibility of the Hybrid Relay Scheme

Correspondingly, the hybrid relay scheme is feasible if the following conditions are fulfilled:

$$\begin{cases} \dfrac{\alpha_{\mathrm{AN}}\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}}-1\right)\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}\eta}}-1\right)M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{L_{\mathrm{B}}\left(N_{\mathrm{t}}-D-N_{\mathrm{r}}\right)\left(N_{\mathrm{r}}-M_{\mathrm{r}}^{\mathrm{B}}\right)} < 1, \\ \dfrac{\alpha_{\mathrm{AN}}\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}}-1\right)\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}(1-\eta)}}-1\right)M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}{L_{\mathrm{B}}\left(N_{\mathrm{r}}-U\right)\left(N_{\mathrm{t}}-M_{\mathrm{t}}^{\mathrm{B}}-N_{\mathrm{r}}\right)} < 1, \\ 0<\eta<1. \end{cases} \tag{6.48}$$

Again, for the proof, refer to [**P7**].

The exact feasibility condition in (6.48) can be simplified by noting that also the first two conditions can in fact be expressed as lower and upper bounds of the duplexing parameter η. Then, the system is feasible as long as there exists a value of η that fulfills

each of these conditions, i.e., the lower bound is less than the upper bound while also being in the open interval $(0, 1)$. In order to derive this feasibility boundary in closed form, it is assumed that

$$\begin{aligned} 2^{\frac{\rho_\mathrm{d}}{\eta}} - 1 &\approx 2^{\frac{\rho_\mathrm{d}}{\eta}}, \\ 2^{\frac{\rho_\mathrm{u}}{1-\eta}} - 1 &\approx 2^{\frac{\rho_\mathrm{u}}{1-\eta}}, \\ 2^{\frac{D\rho_\mathrm{d}}{M_\mathrm{r}^\mathrm{B}\eta}} - 1 &\approx 2^{\frac{D\rho_\mathrm{d}}{M_\mathrm{r}^\mathrm{B}\eta}}, \\ 2^{\frac{U\rho_\mathrm{u}}{M_\mathrm{t}^\mathrm{B}(1-\eta)}} - 1 &\approx 2^{\frac{U\rho_\mathrm{u}}{M_\mathrm{t}^\mathrm{B}(1-\eta)}}. \end{aligned}$$

These approximations make the problem analytically tractable while still resulting in a rather accurate estimate of the feasibility boundary, as will be observed in Section 6.2.4.

Now, with these approximations, the first two conditions can be rewritten as follows:

$$\frac{\rho_\mathrm{d} + \frac{D\rho_\mathrm{d}}{M_\mathrm{r}^\mathrm{B}}}{\log_2\left(\frac{L_\mathrm{B}(N_\mathrm{r}-M_\mathrm{r}^\mathrm{B})(N_\mathrm{t}-D-N_\mathrm{r})}{\alpha_\mathrm{AN} M_\mathrm{r}^\mathrm{B} S_\mathrm{B/d}}\right)} < \eta < 1 - \frac{\rho_\mathrm{u} + \frac{U\rho_\mathrm{u}}{M_\mathrm{t}^\mathrm{B}}}{\log_2\left(\frac{L_\mathrm{B}(N_\mathrm{r}-U)\left(N_\mathrm{t}-M_\mathrm{t}^\mathrm{B}-N_\mathrm{r}\right)}{\alpha_\mathrm{AN} M_\mathrm{t}^\mathrm{B} S_\mathrm{B/u}}\right)}.$$

Moreover, to ensure that the third condition ($0 < \eta < 1$) is fulfilled, it is also required that $\frac{L_\mathrm{B}\left(N_\mathrm{r}-M_\mathrm{r}^\mathrm{B}\right)(N_\mathrm{t}-D-N_\mathrm{r})}{\alpha_\mathrm{AN} M_\mathrm{r}^\mathrm{B} S_\mathrm{B/d}} > 1$ and $\frac{L_\mathrm{B}(N_\mathrm{r}-U)\left(N_\mathrm{t}-M_\mathrm{t}^\mathrm{B}-N_\mathrm{r}\right)}{\alpha_\mathrm{AN} M_\mathrm{t}^\mathrm{B} S_\mathrm{B/u}} > 1$. However, as these conditions are valid with any realistic system parameters, they are not explicitly considered in the continuation. Then, by requiring that the lower bound of η is strictly less than its upper bound, the following approximative feasibility condition for the hybrid relay scheme is obtained:

$$\frac{\rho_\mathrm{d} + \frac{D\rho_\mathrm{d}}{M_\mathrm{r}^\mathrm{B}}}{\log_2\left(\frac{L_\mathrm{B}(N_\mathrm{r}-M_\mathrm{r}^\mathrm{B})(N_\mathrm{t}-D-N_\mathrm{r})}{\alpha_\mathrm{AN} M_\mathrm{r}^\mathrm{B} S_\mathrm{B/d}}\right)} + \frac{\rho_\mathrm{u} + \frac{U\rho_\mathrm{u}}{M_\mathrm{t}^\mathrm{B}}}{\log_2\left(\frac{L_\mathrm{B}(N_\mathrm{r}-U)\left(N_\mathrm{t}-M_\mathrm{t}^\mathrm{B}-N_\mathrm{r}\right)}{\alpha_\mathrm{AN} M_\mathrm{t}^\mathrm{B} S_\mathrm{B/u}}\right)} < 1. \tag{6.49}$$

Then, using (6.49), the highest supported DL data rate requirement of the hybrid relay scheme can be derived in closed form, and it is as follows:

$$\rho_\mathrm{d} < \frac{\log_2\left(\frac{L_\mathrm{B}\left(N_\mathrm{r}-M_\mathrm{r}^\mathrm{B}\right)(N_\mathrm{t}-D-N_\mathrm{r})}{\alpha_\mathrm{AN} M_\mathrm{r}^\mathrm{B} S_\mathrm{B/d}}\right)}{1 + \frac{D}{M_\mathrm{r}^\mathrm{B}}} \left(1 - \frac{\rho_\mathrm{u} + \frac{U\rho_\mathrm{u}}{M_\mathrm{t}^\mathrm{B}}}{\log_2\left(\frac{L_\mathrm{B}(N_\mathrm{r}-U)\left(N_\mathrm{t}-M_\mathrm{t}^\mathrm{B}-N_\mathrm{r}\right)}{\alpha_\mathrm{AN} M_\mathrm{t}^\mathrm{B} S_\mathrm{B/u}}\right)}\right). \tag{6.50}$$

Moreover, the amount of AN SI cancellation that ensures the feasibility of the system is as follows:

$$\alpha_{\mathrm{AN}}^{dB} < 5\log_{10}\left(\frac{L_{\mathrm{B}}^2\left(N_{\mathrm{r}}-U\right)\left(N_{\mathrm{t}}-D-N_{\mathrm{r}}\right)\left(N_{\mathrm{t}}-M_{\mathrm{t}}^{\mathrm{B}}-N_{\mathrm{r}}\right)\left(N_{\mathrm{r}}-M_{\mathrm{r}}^{\mathrm{B}}\right)}{M_{\mathrm{r}}^{\mathrm{B}}M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/d}}S_{\mathrm{B/u}}}\right) - \frac{5}{\log_2\left(10\right)}\left[\rho_{\mathrm{d}}+\rho_{\mathrm{u}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}+\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}+\left[\left(\rho_{\mathrm{d}}-\rho_{\mathrm{u}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}-\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}+\log_2\left(\frac{\left(N_{\mathrm{r}}-U\right)\left(N_{\mathrm{t}}-M_{\mathrm{t}}^{\mathrm{B}}-N_{\mathrm{r}}\right)M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{\left(N_{\mathrm{t}}-D-N_{\mathrm{r}}\right)\left(N_{\mathrm{r}}-M_{\mathrm{r}}^{\mathrm{B}}\right)M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\right)\right)^2+4\left(\rho_{\mathrm{d}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}\right)\left(\rho_{\mathrm{u}}+\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}\right)\right]^{1/2}\right]. \quad (6.51)$$

Again, it should be noted that the above boundary conditions represent the *fundamental* feasibility boundary of the considered system, and they are only reached when the transmit powers tend to infinity. Nevertheless, the fundamental nature of these conditions means that they characterize the ultimate boundary conditions of the proposed self-backhauling solution.

6.2.4 Simulation Results

Next, the proposed system is evaluated with the help of Monte Carlo simulations, considering the three different communications schemes. In the simulations, the UEs are randomly positioned into a circular cell of given size, at the center of which is the self-backhauling AN. The path losses for the different links are then calculated based on the corresponding distances and the adopted path loss model. Moreover, the distribution of the UE locations is uniform, and the DL and UL UEs are allocated from the opposite sides of the cell, as previously mentioned. Note that each UE gets served both in the DL and in the UL by having them alternate between the two modes at regular intervals. By calculating the feasibility boundaries and optimal transmit powers over various random network realizations, the cumulative distribution functions (CDFs) of the corresponding quantities can then be obtained.

Table 6.2 lists all the default system parameters, which are used in the simulations unless otherwise mentioned. Moreover, the path loss model is taken from [201], where a measurement-based model for a center frequency of 3.5 GHz is reported, considering both line-of-sight (LOS) and non-line-of-sight (NLOS) conditions. In the simulations, the LOS model is applied to the link between the AN and the BN, while the NLOS model is used for all the other links. The forthcoming CDFs are obtained by generating 10 000 random UE locations for which the feasibility boundaries or optimal transmit powers are calculated. Furthermore, to ensure a fair comparison between the different schemes, the transmit powers of the half-duplex and hybrid relay schemes are weighted by the proportion of time spent in the corresponding time slot (specified by η), as this more realistically illustrates their overall transmit power usage in relation to the full-duplex scheme.

Table 6.2: The essential default system parameters. Many of the parameter values are also varied in the evaluations.

Parameter	**Value**
Number of AN transmit/receive antennas (N_t / N_r)	200 / 100
Number of DL/UL UEs (D / U)	10 / 10
Number of DL/UL backhaul streams (M_r^B / M_t^B)	12 / 6
Receiver noise floor (σ_n^2)	-90 dBm
Amount of SI cancellation in the AN/BN (α_AN / α_BN)	-120 / -120 dB
Per-UE DL/UL rate requirement (ρ_d / ρ_u)	8 / 2 bps/Hz
Cell radius	50 m
Distance between the AN and the BN	75 m
Number of Monte Carlo simulation runs	10 000

Investigating first the feasibility boundaries, Fig. 6.3 shows the CDFs of the AN SI cancellation requirements in the full-duplex/hybrid relay schemes, using both the exact solutions in (6.45)/(6.48) and the approximated closed-form solutions in (6.47)/(6.51). Firstly, it can be observed that the approximated boundary conditions match very well with the exact values, indicating that the approximative closed-form solutions are indeed rather accurate. It is also evident that the SI cancellation requirements of the full-duplex scheme are less affected by the data rate requirements than those of the hybrid relay scheme. Namely, with the lowest data rate requirements, the full-duplex scheme requires somewhat higher SI cancellation levels in the AN to be feasible, while the opposite is true for the larger data rate requirements. In the former case, the hybrid relay scheme benefits from the fact that it only has to transmit to the UEs *or* to the BN, while the full-duplex scheme must transmit everything at the same time. This somewhat decreases the AN SI cancellation requirements of the hybrid relay scheme, although, with sufficiently high data rate requirements, this benefit is overshadowed by the need to perform duplexing in the time domain.

Considering then the data rate requirements in more detail, Fig. 6.4 shows the CDFs of the highest supported DL data rate requirements for different UL data rate requirements. Again, both the exact and approximative feasibility boundaries are shown, calculated with (6.45)/(6.48) and (6.46)/(6.50), respectively. These results further confirm the accuracy of the approximated boundary conditions, as they are well in line with the exact feasibility boundaries also in this case. Figure 6.4 also indicates that the full-duplex scheme can support higher DL data rates for all the considered UL data rate requirements than the hybrid relay scheme. However, the drawback of the full-duplex scheme is the greater uncertainty in the highest supported DL data rate, evidenced by the lower slope of the CDFs in comparison to the hybrid relay scheme. Hence, while the full-duplex scheme can on average support higher data rates than the hybrid relay scheme, it is more sensitive to the exact network geometry. This suggests that there is a fundamental trade-off between the maximum performance and robustness when comparing the full-duplex and hybrid relay schemes.

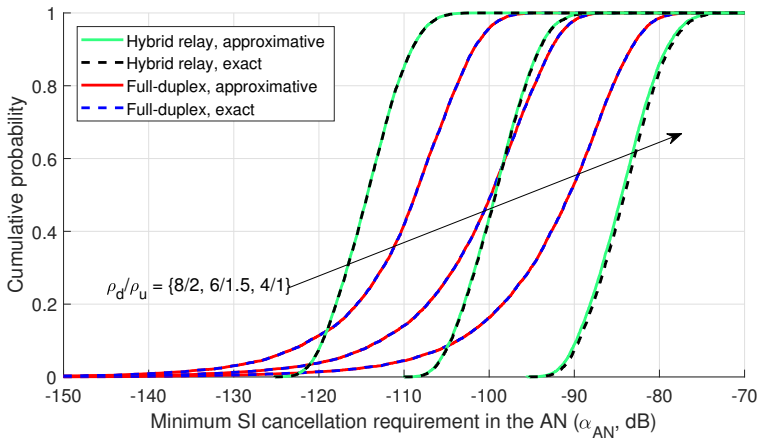


Figure 6.3: CDFs of the minimum AN SI cancellation requirement in the full-duplex and hybrid relay schemes, shown for different DL/UL data rate requirements.

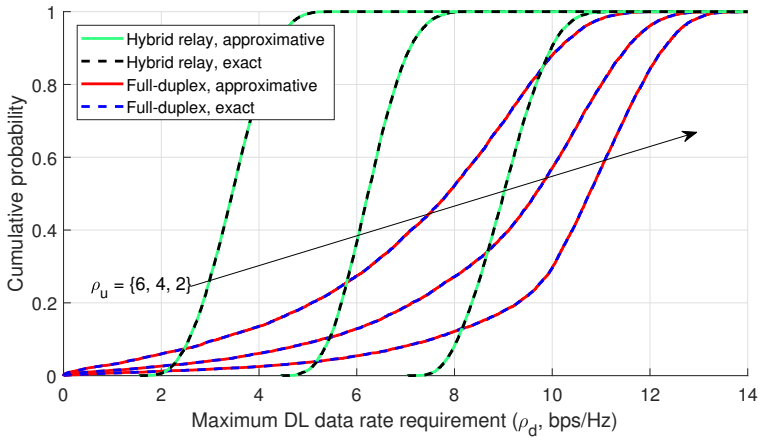


Figure 6.4: CDFs of the maximum supported DL data rate requirement in the full-duplex and hybrid relay schemes, shown for different UL data rate requirements.

Then, analyzing the transmit power efficiency of the different schemes, the CDFs of the optimal transmit powers for the different communicating parties are shown in Fig. 6.5, using the default system parameters. Firstly, it can be observed that the full-duplex scheme is capable of operating with the lowest transmit powers. However, the drawback of the full-duplex scheme is its inability to achieve the QoS requirements under all network geometries, evidenced by the fact that the CDFs converge to a value less than 1. These infeasible system scenarios are the ones for which the conditions in (6.45) are not fulfilled and they are also evident in Fig. 6.3, which indicates that the feasibility of the full-duplex scheme often requires SI cancellation levels beyond -120 dB in the AN when $\rho_{\mathrm{d}} = 8$ bps/Hz and $\rho_{\mathrm{u}} = 2$ bps/Hz.

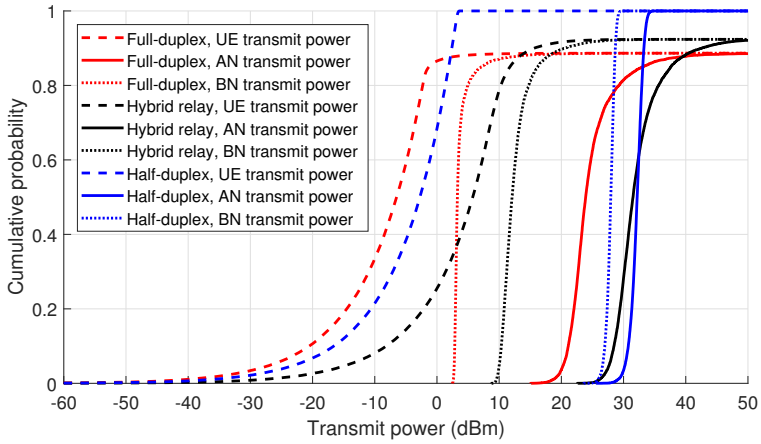


Figure 6.5: CDFs of the transmit powers of the individual communicating parties with the default system parameters.

The hybrid relay scheme is the next best option when considering the overall transmit power usage, although it is outperformed by the half-duplex scheme in terms of the UE transmit power consumption. The latter is caused by the division of the time slots in the two schemes, which require the half-duplex scheme to backhaul the DL data simultaneously with the UL transmissions. As the DL data rate requirements are higher, this naturally increases the length of the corresponding time slot, allowing for a lower UE transmit power. This phenomenon is simply a result of the formulation of the optimization problem, which aims at minimizing the *total* transmit power consumption. A different outcome would be obtained if more weight was given to minimizing the UE transmit powers in all the communications schemes. Furthermore, another drawback of the hybrid relay scheme is its inability to fulfill the QoS requirements under some network geometries, similar to the full-duplex scheme.

Then, finally, Fig. 6.6 shows the CDFs of the *total* transmit power consumption in the different communications schemes, using three values of α_{AN}. Firstly, it can be seen that the full-duplex scheme is in general capable of obtaining the QoS requirements with the lowest transmit power usage, regardless of the AN SI cancellation performance. However, again, it suffers from being infeasible under certain network configurations if there is residual SI in the AN, meaning that it cannot fulfill the QoS requirements in these cases with any finite transmit powers. With perfect SI cancellation in the AN, there are no infeasible network configurations, although rather high transmit powers are still required in some extreme cases, as evidenced by Fig. 6.6.

Investigating the transmit power consumption of the other communications schemes, it can be observed that, with the higher AN SI cancellation levels, the hybrid relay scheme outperforms the half-duplex scheme. This is particularly evident with perfect SI cancellation, although already -120 dB of SI suppression is in most cases enough for the hybrid relay scheme to provide lower transmit powers than the half-duplex scheme. However, with imperfect SI cancellation in the AN, there are some network configurations which are infeasible for the hybrid relay scheme, meaning that it cannot obtain the

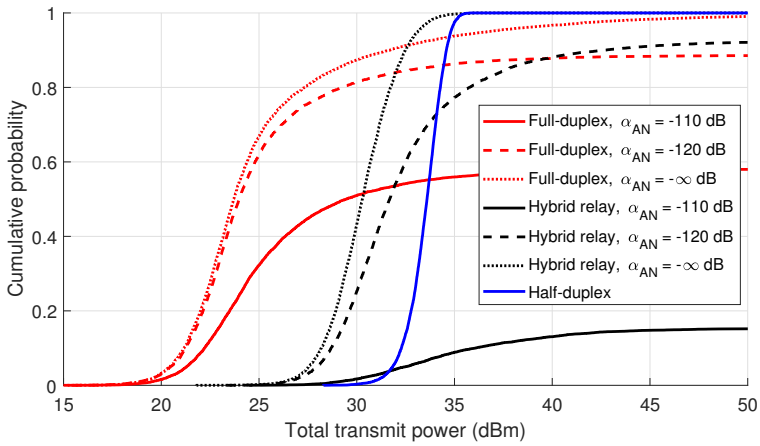


Figure 6.6: CDFs of the total used transmit power in each scheme, shown for different amounts of SI cancellation.

required data rates with any finite transmit powers. Especially, when $\alpha_{\mathrm{AN}} = -110$ dB, the hybrid relay scheme can fulfill the QoS requirements only for roughly 15% of the random network configurations, indicating that this scheme requires rather high SI cancellation levels in the AN to perform reliably. The same conclusions can be drawn also from Fig. 6.3, which clearly shows that the AN SI cancellation requirement in the hybrid relay scheme is more than -110 dB for a large portion of the network geometries when considering the rate requirements $\rho_{\mathrm{d}} = 8$ bps/Hz and $\rho_{\mathrm{u}} = 2$ bps/Hz.

In conclusion, it can be said that in most cases the full-duplex scheme provides the highest transmit power efficiency when operating under some given QoS requirements. However, care must be taken to ensure that the system parameters are such that the full-duplex scheme does not suffer from the problem of infeasibility. This means, for instance, that the AN SI cancellation performance must be sufficiently high with respect to the data rate requirements. If, however, this cannot be ensured, then the hybrid relay scheme is perhaps the more preferable option as it is somewhat more robust in obtaining the given data rate requirements under different circumstances. Alternatively, the system could also switch back and forth between the full-duplex, hybrid relay, and half-duplex schemes, depending on which of them is the best option for the prevailing network geometry. Nevertheless, the analysis and the findings indicate that, altogether, the IBFD capability is a greatly beneficial feature for the considered self-backhauling AN.

CHAPTER 7

Summary

This final chapter provides a summary of the most significant results and findings of this thesis while also laying out possible future research directions. Namely, considering the fast pace of research into IBFD radios, there is little doubt that they will play a role also in the wireless communications systems of the future. Therefore, many highly relevant research items still remain, regardless of the various results already reported in this thesis.

7.1 Main Results

To provide some initial understanding into the challenges related to digital SI cancellation in IBFD radios, Chapter 3 and publications [**P1, P2**] concentrated on the effects of the different analog imperfections. In particular, their impact on the overall residual SI waveform was determined by deriving their power levels after the analog-to-digital conversion in the RX chain. The hereby obtained findings suggest that the most significant analog impairments in the context of IBFD transceivers are the I/Q imbalance and the nonlinear distortion produced by the TX PA. These observations set the foundation for the digital cancellation solutions and the signal models utilized therein.

Then, Chapter 4 and publications [**P3**–**P6**] presented four different signal models for digital SI cancellation, alongside with two different parameter estimation solutions. One of the signal models was considered as the benchmark solutions as it does not consider any of the impairments, while the three other models incorporate the PA-induced nonlinearity, I/Q imbalance, or both of them. The two original parameter estimation schemes are based on LS and LMS, respectively, the former being a block-wise procedure while the latter is an iterative algorithm for adaptively estimating the SI channel coefficients. Moreover, an original complexity reduction scheme was also proposed in this chapter, which utilizes PCA to reduce the number of coefficients that must be estimated.

The proposed cancellation solutions were then evaluated in Chapter 5, using the simulation and measurement results reported in [**P1**–**P6**]. First, the different digital cancellers were compared with the help of comprehensive waveform simulations modeling a MIMO IBFD transceiver under realistic impairment levels. With both PA-induced nonlinearities and I/Q imbalance present in the transceiver, the results suggest that

both of these impairments must be included in the signal models to obtain sufficient SI cancellation performance. Furthermore, the complexity reduction scheme was also shown to be capable of reducing the number of coefficients by as much as 65% without any decrease in the cancellation accuracy.

The measurement results shown in Chapter 5 further confirmed the high performance of the proposed nonlinear digital canceller under real-life conditions. When using a shared TX/RX antenna and an active RF canceller, the nonlinear digital canceller could suppress the SI to the level of the receiver noise floor with bandwidths ranging from 20 MHz to 80 MHz. With the used transmit power levels, this translates to roughly 90 dB of SI suppression. Moreover, due to the low-cost PA, a linear digital canceller could not fully suppress the residual SI in the digital domain, indicating that nonlinear modeling is indeed necessary in a realistic system. The performance of the nonlinear digital canceller was also successfully evaluated in an IBFD relay with a high-isolation back-to-back antenna. There, even when using only a digital canceller without any active cancellation in the analog domain, the SI was cancelled to the level of the receiver noise floor, again for bandwidths between 20–80 MHz. This means that the total amount of SI suppression, obtained with only physical antenna isolation and active digital cancellation, was in the order of 100–110 dB. These are some of the highest reported overall SI cancellation performances to date.

As the last and latest contribution of this thesis, Chapter 6 and publication **[P7]** analyzed a potential use-case for an IBFD-capable transceiver: a self-backhauling AN that uses the same time-frequency resource for serving the UEs and for backhauling the DL and UL data wirelessly. The greatest challenge for such a system is managing the various interference sources, such as the UL-to-DL IUI, SI, and the interference between the BN and the UEs. This thesis proposed a solution where the interference is managed by a QoS-fulfilling power allocation scheme that minimizes the transmit powers of all communicating parties while still ensuring that the required data rates are obtained. Moreover, the purely full-duplex AN was compared to two reference schemes where some of the transmissions are performed in separate time slots. The optimal power allocation was derived for all the considered schemes in closed form, alongside with the fundamental feasibility conditions, the latter of which define whether the given data rates can even be obtained in the first place. The numerical results showed that the purely full-duplex AN can typically obtain the QoS requirements with the lowest transmit powers, although it might be entirely unfeasible for certain network geometries.

Altogether, this thesis demonstrates that, by careful modeling of the residual SI in the digital domain, true wireless IBFD communications is indeed possible. This requires advanced signal models that incorporate also some of the analog imperfections into the cancellation processing, the benefit of this increased modeling complexity being the near-perfect suppression of the SI in the digital domain. The high cancellation performance of the proposed digital cancellers was confirmed both by extensive simulations and real-world measurements, the latter being performed for two different IBFD prototypes. Both of the considered prototypes could cancel the SI to the level of the receiver noise floor, thereby presenting important guidelines for what is required when implementing IBFD-capable devices in practice. Moreover, significant gains can also be obtained by utilizing IBFD-capable transceivers on a system level, meaning that IBFD radios will likely play an important role in the development of the future wireless networks.

7.2 Future Work

The promising results of this thesis demonstrating the practical feasibility of wireless IBFD communications open various avenues also for further research. Firstly, and perhaps most importantly, the network-level throughput gains provided by IBFD-capable radio devices need to be experimentally demonstrated. Namely, most existing experimental works show the data rate improvement only for individual and isolated links, if at all, which is not entirely representative of a practical deployment scenario. Investigating a network consisting of several nodes, containing one or several IBFD-capable transceivers, will reveal the full potential of the IBFD technology.

Furthermore, performing efficient SI cancellation with transmit powers beyond 30 dBm is another area that still requires further research. Especially, the results reported this thesis are achieved using transmit powers of 24 dBm or lower, which is insufficient for applications that require a wide area of coverage, such as a macro cell with an IBFD-capable base station. Employing such high transmit powers is likely to require even more advanced signal models for sufficiently accurate SI regeneration, which also emphasizes the role of the different complexity reduction schemes to maintain the computational requirements of the digital cancellation procedure on a reasonable level. Likewise, extending the SI cancellation architectures to support the lower frequencies commonly used in military communications systems is also a highly relevant topic for future research.

In addition, exploring the idea of IBFD-capable large-array ANs further, both theoretically and experimentally, is yet another potential future research direction. The system considered in this thesis is only one of the many possibilities for utilizing such ANs, and hence several potential research topics still remain. Moreover, experimentally demonstrating true IBFD operation with large TX and RX antenna arrays is something that has received relatively little attention.

In summary, while significant progress has already been made in this field of research, various open research topics still remain. Investigating these aspects further will likely push the wireless IBFD technology closer to the level of maturity that is needed for commercially viable devices.

References

[1] A. Abdelhafiz, A. Kwan, O. Hammi, and F. M. Ghannouchi, "Digital predistortion of LTE-A power amplifiers using compressed-sampling-based unstructured pruning of Volterra series," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 11, pp. 2583–2593, Nov. 2014.

[2] A. A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1399–1410, Dec. 1995.

[3] V. Aggarwal, M. Duarte, A. Sabharwal, and N. Shankaranarayanan, "Full- or half-duplex? A capacity analysis with bounded radio resources," in *Proc. Information Theory Workshop (ITW)*, Sep. 2012, pp. 207–211.

[4] M. AghababaeeTafreshi, M. Koskela, D. Korpi, P. Jääskeläinen, M. Valkama, and J. Takala, "Software defined radio implementation of adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio," in *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec. 2016, pp. 733–737.

[5] E. Ahmed, A. Eltawil, and A. Sabharwal, "Simultaneous transmit and sense for cognitive radios using full-duplex: A first study," in *Proc. IEEE International Symposium on Antennas and Propagation*, Jul. 2012.

[6] E. Ahmed and A. M. Eltawil, "All-digital self-interference cancellation technique for full-duplex systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3519–3532, Jul. 2015.

[7] E. Ahmed and A. M. Eltawil, "On phase noise suppression in full-duplex systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1237–1251, Mar. 2015.

[8] E. Ahmed, A. M. Eltawil, Z. Li, and B. A. Cetiner, "Full-duplex systems using multireconfigurable antennas," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 5971–5983, Nov. 2015.

[9] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Rate gain region and design tradeoffs for full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3556–3565, Jul. 2013.

[10] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Self-interference cancellation with nonlinear distortion suppression for full-duplex systems," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013, pp. 1199–1203.

[11] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Self-interference cancellation with phase noise induced ICI suppression for full-duplex systems," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013, pp. 3384–3388.

[12] M. O. Al-Kadri, A. Aijaz, and A. Nallanathan, "An energy-efficient full-duplex MAC protocol for distributed wireless networks," *IEEE Wireless Communications Letters*, vol. 5, no. 1, pp. 44–47, Feb. 2016.

[13] M. S. Amjad and O. Gurbuz, "Linear digital cancellation with reduced computational complexity for full-duplex radios," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2017.

[14] C. R. Anderson, S. Krishnamoorthy, C. G. Ranson, T. J. Lemon, W. G. Newhall, T. Kummetz, and J. H. Reed, "Antenna isolation, wideband multipath propagation measurements, and interference mitigation for on-frequency repeaters," in *Proc. IEEE SoutheastCon*, Mar. 2004, pp. 110–114.

[15] L. Anttila, P. Händel, and M. Valkama, "Joint mitigation of power amplifier and I/Q modulator impairments in broadband direct-conversion transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 4, pp. 730–739, Apr. 2010.

[16] L. Anttila, M. Valkama, and M. Renfors, "Circularity-based I/Q imbalance compensation in wideband direct-conversion receivers," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2099–2113, Jul. 2008.

[17] L. Anttila, D. Korpi, E. Antonio-Rodríguez, R. Wichman, and M. Valkama, "Modeling and efficient cancellation of nonlinear self-interference in MIMO full-duplex transceivers," in *Proc. IEEE Globecom Workshops*, Dec. 2014, pp. 862–868.

[18] L. Anttila, D. Korpi, V. Syrjälä, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013, pp. 1193–1198.

[19] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, "MIDU: Enabling MIMO full duplex," in *Proc. 18th Annual International Conference on Mobile Computing and Networking (MobiCom)*, Aug. 2012, pp. 257–268.

[20] R. Askar, T. Kaiser, B. Schubert, T. Haustein, and W. Keusgen, "Active self-interference cancellation mechanism for full-duplex wireless transceivers," in *Proc. Ninth International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Jun. 2014, pp. 539–544.

[21] R. Askar, B. Schubert, W. Keusgen, and T. Haustein, "Full-duplex wireless transceiver in presence of I/Q mismatches: Experimentation and estimation algorithm," in *Proc. IEEE Globecom Workshops*, Dec. 2015.

[22] A. C. M. Austin, A. Balatsoukas-Stimming, and A. Burg, "Digital predistortion of power amplifier non-linearities for full-duplex transceivers," in *Proc. 17th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2016.

[23] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[24] J. Bai and A. Sabharwal, "Asymptotic analysis of MIMO multi-cell full-duplex networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 4, pp. 2168–2180, Apr. 2017.

[25] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 2, pp. 1100–1114, Feb. 2015.

[26] A. Balatsoukas-Stimming, A. Austin, P. Belanovic, and A. Burg, "Baseband and RF hardware impairments in full-duplex wireless systems: Experimental characterisation and suppression," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 142, May 2015.

[27] P. D. L. Beasley, A. G. Stove, B. J. Reits, and B. As, "Solving the problems of a single antenna frequency modulated CW radar," in *Proc. IEEE International Radar Conference*, May 1990, pp. 391–395.

[28] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. SIGCOMM'13*, Aug. 2013, pp. 375–386.

[29] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proc. 11th USENIX Conference on Networked Systems Design and Implementation (NSDI)*, Apr. 2014, pp. 359–372.

[30] S. Bigelow, J. Carr, and S. Winder, *Understanding Telephone Electronics*, 4th ed. Newnes, 2001.

[31] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[32] D. W. Bliss, "Full-duplex self-interference mitigation analysis for direct conversion RF nonlinear MIMO channel models with IQ mismatch," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017, pp. 6543–6547.

[33] D. W. Bliss and K. W. Forsythe, "MIMO radar medical imaging: Self-interference mitigation for breast tumor detection," in *Proc. 40th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Oct. 2006, pp. 1558–1562.

[34] D. W. Bliss, T. M. Hancock, and P. Schniter, "Hardware phenomenological effects on cochannel full-duplex MIMO relay performance," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2012, pp. 34–39.

[35] D. W. Bliss, P. A. Parker, and A. R. Margetts, "Simultaneous transmission and reception for improved wireless network performance," in *Proc. 14th IEEE/SP Workshop on Statistical Signal Processing (SSP)*, Aug. 2007, pp. 478–482.

[36] D. W. Bliss and Y. Rong, "Effects of channel estimation errors on in-band full-duplex MIMO radios using adaptive transmit spatial mitigation," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013, pp. 9–13.

[37] D. W. Bliss and Y. Rong, "Full-duplex self-interference mitigation performance in nonlinear channels," in *Proc. 48th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2014, pp. 1696–1700.

[38] T. C. Carusone, D. A. Johns, and K. W. Martin, *Analog Integrated Circuit Design.* Wiley, 1997.

[39] J. K. Cavers and M. W. Liao, "Adaptive compensation for imbalance and offset losses in direct conversion transceivers," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 581–588, Nov. 1993.

[40] B. K. Chalise, H. A. Suraweera, G. Zheng, and G. K. Karagiannidis, "Beamforming optimization for full-duplex wireless-powered MIMO systems," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3750–3764, Sep. 2017.

[41] G. Chen, Y. Gong, P. Xiao, and R. Tafazolli, "Dual antenna selection in self-backhauling multiple small cell networks," *IEEE Communications Letters*, vol. 20, no. 8, pp. 1611–1614, Aug. 2016.

[42] L. Chen, F. R. Yu, H. Ji, V. C. M. Leung, X. Li, and B. Rong, "A full-duplex self-backhaul scheme for small cell networks with massive MIMO," in *Proc. IEEE International Conference on Communications (ICC)*, May 2016.

[43] S. Y. Chen, T. F. Huang, K. C. J. Lin, Y. W. P. Hong, and A. Sabharwal, "Probabilistic-based adaptive full-duplex and half-duplex medium access control," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2015.

[44] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel full duplex wireless communication," in *Proc. 16th Annual International Conference on Mobile Computing and Networking (MobiCom)*, Sep. 2010, pp. 1–12.

[45] W. Choi and H. Lim, "Immediate acknowledgement for single-channel full-duplex wireless networks," in *Proc. 9th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS)*, Oct. 2012, pp. 477–478.

[46] W. Choi, H. Lim, and A. Sabharwal, "Power-controlled medium access control protocol for full-duplex WiFi networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3601–3613, Jul. 2015.

[47] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.

[48] B. Chun, E.-R. Jeong, J. Joung, Y. Oh, and Y. H. Lee, "Pre-nulling for self-interference suppression in full-duplex relays," in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Oct. 2009, pp. 91–97.

[49] B. Chun and Y. H. Lee, "A spatial self-interference nullification method for full duplex amplify-and-forward MIMO relays," in *Proc. IEEE Wireless Communication and Networking Conference*, Apr. 2010.

[50] M. Chung, M. S. Sim, J. Kim, D. K. Kim, and C. b. Chae, "Prototyping real-time full duplex radios," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 56–63, Sep. 2015.

[51] W. Chung, D. Hong, R. Wichman, and T. Riihonen, "Interference cancellation architecture for full-duplex system with GFDM signaling," in *Proc. 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 788–792.

[52] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. MIT Press, 2009.

[53] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3702–3713, Jul. 2012.

[54] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1541–1553, Sep. 2012.

[55] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[56] B. Debaillie, D. J. van den Broek, C. Lavín, B. van Liempd, E. A. M. Klumperink, C. Palacios, J. Craninckx, B. Nauta, and A. Pärssinen, "Analog/RF solutions enabling compact full-duplex radios," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1662–1673, Sep. 2014.

[57] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[58] R. Doost-Mohammady, M. Y. Naderi, and K. R. Chowdhury, "Performance analysis of CSMA/CA based medium access in full duplex wireless communications," *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1457–1470, Jun. 2016.

[59] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[60] M. Duarte, A. Feki, and S. Valentin, "Inter-user interference coordination in full-duplex systems based on geographical context information," in *Proc. IEEE International Conference on Communications (ICC)*, May 2016.

[61] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. Ramakrishnan, C. Rice, and N. Shankaranarayanan, "Design and characterization of a full-duplex multiantenna system for WiFi networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 3, pp. 1160–1177, Mar. 2014.

[62] M. Duarte, "Full-duplex wireless: Design, implementation and characterization," Ph.D. dissertation, Rice University, 2012.

[63] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. 44th Asilomar Conference on Signals, Systems, and Computers (ASILOMAR)*, Nov. 2010, pp. 1558–1562.

[64] R. Durrett, *Probability: Theory and Examples*, 4th ed. Cambridge University Press, 2010.

[65] A. El-Keyi and H. Yanikomeroglu, "Interference alignment for heterogeneous full-duplex cellular networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2016.

[66] M. Emara, M. Faerber, L. G. Baltar, J. Nossek, and K. Roth, "Nonlinear digital self-interference cancellation with reduced complexity for full duplex systems," in *Proc. International ITG Workshop on Smart Antennas (WSA)*, Mar. 2017.

[67] Ericsson AB, "5G radio access," white paper, 2015. [Online]. Available: https://www.ericsson.com/assets/local/publications/white-papers/wp-5g.pdf

[68] "LTE; evolved universal terrestrial radio access (E-UTRA); base station (BS) radio transmission and reception (3GPP TS 36.104, version 14.3.0, release 14)," ETSI, Sophia Antipolis Cedex, France, Mar. 2017.

[69] "LTE; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception (3GPP TS 36.101 version 14.3.0 release 14)," ETSI, Sophia Antipolis Cedex, France, Jan. 2017.

[70] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *Proc. 45th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2011, pp. 2002–2006.

[71] E. Everett, C. Shepard, L. Zhong, and A. Sabharwal, "Softnull: Many-antenna full-duplex wireless via digital beamforming," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8077–8092, Dec. 2016.

[72] E. Everett and A. Sabharwal, "Spatial degrees-of-freedom in large-array full-duplex: The impact of backscattering," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, Dec. 2016.

[73] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[74] P. Ferrand and M. Duarte, "Multi-tap digital canceller for full-duplex applications," in *Proc. 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2017.

[75] F. M. Ghannouchi, "Power amplifier and transmitter architectures for software defined radio systems," *IEEE Circuits and Systems Magazine*, vol. 10, no. 4, pp. 56–63, Fourth Quarter 2010.

[76] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microwave Magazine*, vol. 10, no. 7, pp. 52–64, Dec. 2009.

[77] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, Jun. 2014.

[78] S. Goyal, P. Liu, O. Gurbuz, E. Erkip, and S. Panwar, "A distributed MAC protocol for full duplex radio," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013, pp. 788–792.

[79] Q. Gu, *RF System Design of Transceivers for Wireless Communications*. Springer, 2006.

[80] H. Hamazumi, K. Imamura, N. Iai, K. Shibuya, and M. Sasaki, "A study of a loop interference canceller for the relay stations in an SFN for digital terrestrial broadcasting," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, Dec. 2000, pp. 167–171 vol.1.

[81] O. Hammi, A. Abdelhafiz, F. M. Ghannouchi, and T. Y. Al-Naffouri, "On the use of compressed sampling algorithms for impairments compensation in dynamic nonlinear transmitters," in *Proc. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Dec. 2015, pp. 641–645.

[82] K. Haneda, M. Valkama, T. Riihonen, E. Antonio-Rodríguez, and D. Korpi, "Design and implementation of full-duplex transceivers," in *Signal Processing for 5G: Algorithms and Implementations*, F.-L. Luo and C. J. Zhang, Eds. Wiley, 2016, ch. 17.

[83] I. Harjula, R. Wichman, K. Pajukoski, E. Lähetkangas, E. Tiirola, and O. Tirkkonen, "Full duplex relaying for local area," in *Proc. 24th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2013, pp. 2684–2688.

[84] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Prentice Hall, 1996.

[85] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Prentice Hall, 1999.

[86] A. Hazmi, J. Rinne, and M. Renfors, "Cancellation of loop interference with exponential profile using autocorrelation method in OFDM based systems," in *Proc. Ninth International Conference on Communications Systems (ICCS)*, Sep. 2004, pp. 140–144.

[87] A. Hazmi, J. Rinne, and M. Renfors, "Enhanced LMS based feedback loop interference cancellation in DVB-H systems with gap fillers," in *Proc. IEEE International Conference on Signal Processing and Communications (ICSPC)*, Nov. 2007, pp. 943–946.

[88] M. Heino, S. Venkatasubramanian, C. Icheln, and K. Haneda, "Design of wavetraps for isolation improvement in compact in-band full-duplex relay antennas," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 3, pp. 1061–1070, Mar. 2016.

[89] "HMC631LP3 / 631LP3E, GaAs HBT vector modulator 1.8-2.7 GHz," Hittite Microwave Corporation, Chelmsford, Massachusetts, USA.

[90] J. K. Hong, Y. W. Suh, J. Y. Choi, and J. S. Seo, "Echo canceller for on-channel repeaters in T-DMB system," in *Proc. 10th International Conference on Advanced Communication Technology (ICACT)*, vol. 3, Feb. 2008, pp. 1735–1738.

[91] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114–121, Feb. 2014.

[92] M. Horowitz, "Computing's energy problem (and what we can do about it)," in *Proc. IEEE International Solid-State Circuits Conference (ISSCC)*, Feb. 2014, pp. 10–14.

[93] H. T. Howard and G. L. Tyler, "Bistatic-radar investigation," in *Apollo 14 Preliminary Science Report*. National Aeronautics and Space Administration, 1971, ch. 17, NASA SP-272.

[94] H. T. Howard and G. L. Tyler, "Bistatic-radar investigation," in *Apollo 15 Preliminary Science Report*. National Aeronautics and Space Administration, 1972, ch. 23, NASA SP-289.

[95] H. T. Howard and G. L. Tyler, "Bistatic-radar investigation," in *Apollo 16 Preliminary Science Report*. National Aeronautics and Space Administration, 1972, ch. 25, NASA SP-315.

[96] Y. Hua, Y. Ma, A. Gholian, Y. Li, A. C. Cirik, and P. Liang, "Radio self-interference cancellation by transmit beamforming, all-analog cancellation and blind digital tuning," *Signal Processing*, vol. 108, no. 2015, pp. 322–340, Oct. 2014.

[97] X. Huang, K. Yang, F. Wu, and S. Leng, "Power control for full-duplex relay-enhanced cellular networks with QoS guarantees," *IEEE Access*, vol. 5, pp. 4859–4869, Mar. 2017.

[98] Huawei Technologies Co. Ltd., "5G: A technology vision," white paper, 2013. [Online]. Available: http://www.huawei.com/ilink/en/download/HW_314849

[99] T. Huusari, Y.-S. Choi, P. Liikkanen, D. Korpi, S. Talwar, and M. Valkama, "Wideband self-adaptive RF cancellation circuit for full-duplex radio: Operating principle and measurements," in *Proc. 81st IEEE Vehicular Technology Conference (VTC Spring)*, May 2015.

[100] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Wiley, 2001.

[101] M. Isaksson, D. Wisell, and D. Rönnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 1, pp. 348–359, Jan. 2006.

[102] "ITU-T recommendation G.992.1: Asymmetric digital subscriber line (ADSL) transceivers," ITU, Geneva, Switzerland, Jun. 1999.

[103] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. 17th Annual International Conference on Mobile computing and Networking (MobiCom)*, Sep. 2011, pp. 301–312.

[104] G. James, D. Burley, D. Clements, P. Dyke, J. Searl, N. Steele, and J. Wright, *Advanced Modern Engineering Mathematics*, 4th ed. Prentice Hall, 2011.

[105] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," *IEEE Transactions on Broadcasting*, vol. 54, no. 2, pp. 257–268, Jun. 2008.

[106] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 3933–3938, Aug. 2009.

[107] Y. Y. Kang and J. H. Cho, "Capacity of MIMO wireless channel with full-duplex amplify-and-forward relay," in *Proc. 20th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2009, pp. 117–121.

[108] Y. Y. Kang, B. J. Kwak, and J. H. Cho, "An optimal full-duplex AF relay for joint analog and digital domain self-interference cancellation," *IEEE Transactions on Communications*, vol. 62, no. 8, pp. 2758–2772, Aug. 2014.

[109] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Prentice Hall, 1993.

[110] P. Kela, J. Turkka, and M. Costa, "Borderless mobility in 5G outdoor ultra-dense networks," *IEEE Access*, vol. 3, pp. 1462–1476, Aug. 2015.

[111] J. L. Kelly, "Self-adaptive echo canceller," U.S. Patent 3 500 000, Mar. 10, 1970.

[112] C. Y. Kim, J. G. Kim, and S. Hong, "A quadrature radar topology with Tx leakage canceller for 24-GHz radar applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 55, no. 7, pp. 1438–1444, Jul. 2007.

[113] J.-G. Kim, S. Ko, S. Jeon, J.-W. Park, and S. Hong, "Balanced topology to cancel Tx leakage in CW radar," *IEEE Microwave and Wireless Components Letters*, vol. 14, no. 9, pp. 443–445, Sep. 2004.

[114] K. Kim, S. W. Jeon, and D. K. Kim, "The feasibility of interference alignment for full-duplex MIMO cellular networks," *IEEE Communications Letters*, vol. 19, no. 9, pp. 1500–1503, Sep. 2015.

[115] S. J. Kim, J. Y. Lee, J. C. Lee, J. H. Kim, B. Lee, and N. Y. Kim, "Adaptive feedback interference cancellation system (AF-ICS)," in *Proc. IEEE MTT-S International Microwave Symposium (IMS)*, Jun. 2003, pp. 627–630.

[116] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1150–1158, Nov. 2003.

[117] M. E. Knox, "Single antenna full duplex communications using a common carrier," in *Proc. 13th Annual IEEE Wireless and Microwave Technology Conference (WAMICON)*, Apr. 2012.

[118] D. Knuth, *The Art of Computer Programming*, 3rd ed. Addison–Wesley, 1997, vol. 1, Fundamental Algorithms.

[119] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap RF canceller for in-band full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4321–4334, Jun. 2016.

[120] D. Korpi, M. AghababaeeTafreshi, M. Piililä, L. Anttila, and M. Valkama, "Advanced architectures for self-interference cancellation in full-duplex radios: Algorithms and measurements," in *Proc. 50th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2016, pp. 1553–1557.

[**121**] D. Korpi, L. Anttila, and M. Valkama, "Feasibility of in-band full-duplex radio transceivers with imperfect RF components: analysis and enhanced cancellation algorithms," in *Proc. Ninth International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM)*, Jul. 2014, pp. 532–538.

[**122**] D. Korpi, L. Anttila, and M. Valkama, "Impact of received signal on self-interference channel estimation and achievable rates in in-band full-duplex transceivers," in *Proc. 48th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2014, pp. 975–982.

[**123**] D. Korpi, L. Anttila, and M. Valkama, "Reference receiver based digital self-interference cancellation in MIMO full-duplex transceivers," in *Proc. IEEE Globecom Workshops*, Dec. 2014, pp. 1001–1007.

[**124**] D. Korpi, L. Anttila, and M. Valkama, "Asymmetric full-duplex with contiguous downlink carrier aggregation," in *Proc. 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2016.

[**125**] D. Korpi, Y.-S. Choi, T. Huusari, S. Anttila, L. Talwar, and M. Valkama, "Adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio: algorithms and RF measurements," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2015.

[**126**] D. Korpi, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Digital self-interference cancellation under nonideal RF components: Advanced algorithms and measured performance," in *Proc. 16th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2015, pp. 286–290.

[**127**] D. Korpi, T. Riihonen, K. Haneda, K. Yamamoto, and M. Valkama, "Achievable transmission rates and self-interference channel estimation in hybrid full-duplex/half-duplex MIMO relaying," in *Proc. 82nd IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2015.

[**128**] D. Korpi, T. Riihonen, and M. Valkama, "Achievable rate regions and self-interference channel estimation in hybrid full-duplex/half-duplex radio links," in *Proc. 49th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2015.

[**129**] D. Korpi, T. Riihonen, and M. Valkama, "Self-backhauling full-duplex access node with massive antenna arrays: Power allocation and achievable sum-rate," in *Proc. 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 1618–1622.

[**130**] D. Korpi, T. Riihonen, and M. Valkama, "Feasibility of self-backhauling in full-duplex radio access systems under QoS constraints," in *Proc. IEEE International Conference on Communications (ICC)*, May 2017, pp. 749–754.

[**131**] D. Korpi, T. Riihonen, and M. Valkama, "Inband full-duplex radio access system with self-backhauling: Transmit power minimization under QoS requirements," in

Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mar. 2017, pp. 6558–6562.

[132] D. Korpi, M. Valkama, T. Riihonen, and R. Wichman, "Implementation challenges in full-duplex radio transceiver," in *Proc. XXXIII Finnish URSI Convention on Radio Science*, Apr. 2013, pp. 181–184.

[133] D. Korpi, S. Venkatasubramanian, T. Riihonen, L. Anttila, S. Otewa, C. Icheln, K. Haneda, S. Tretyakov, M. Valkama, and R. Wichman, "Advanced self-interference cancellation and multiantenna techniques for full-duplex radios," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013, pp. 3–8.

[134] D. Korpi, T. Riihonen, A. Sabharwal, and M. Valkama, "Sum-rate analysis and optimization of self-backhauling based full-duplex radio access system," *arXiv Online Library*, 2017, unpublished. [Online]. Available: http://arxiv.org/abs/1604.06571

[135] I. Krikidis and H. A. Suraweera, "Full-duplex cooperative diversity with Alamouti space-time code," *IEEE Wireless Communications Letters*, vol. 2, no. 5, pp. 519–522, Oct. 2013.

[136] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4381–4393, Dec. 2012.

[137] I. Krikidis, H. A. Suraweera, S. Yang, and K. Berberidis, "Full-duplex relaying over block fading channel: A diversity perspective," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4524–4535, Dec. 2012.

[138] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[139] P. Larsson and M. Prytz, "MIMO on-frequency repeater with self-interference cancellation and mitigation," in *Proc. 69th IEEE Vehicular Technology Conference (VTC Spring)*, Apr. 2009.

[140] L. Laughlin, C. Zhang, M. A. Beach, K. A. Morris, and J. Haine, "A widely tunable full duplex transceiver combining electrical balance isolation and active analog cancellation," in *Proc. 81st IEEE Vehicular Technology Conference (VTC Spring)*, May 2015.

[141] L. Laughlin, C. Zhang, M. A. Beach, K. A. Morris, and J. L. Haine, "Dynamic performance of electrical balance duplexing in a vehicular scenario," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 844–847, 2017.

[142] J. H. Lee, "Self-interference cancelation using phase rotation in full-duplex wireless," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4421–4429, Nov. 2013.

[143] R. Li, A. Masmoudi, and T. Le-Ngoc, "Self-interference cancellation with phase-noise suppression in full-duplex systems," in *Proc. 26th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Aug. 2015, pp. 261–265.

[144] S. Li and R. D. Murch, "An investigation into baseband techniques for single-channel full-duplex wireless communication systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 9, pp. 4794–4806, Sep. 2014.

[145] S. Li and R. Murch, "Full-duplex wireless communication using transmitter output based echo cancellation," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, Dec. 2011.

[146] Z. Li, E. Ahmed, A. M. Eltawil, and B. A. Cetiner, "A beam-steering reconfigurable antenna for WLAN applications," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 1, pp. 24–32, Jan. 2015.

[147] Y. Liao, K. Bian, L. Song, and Z. Han, "Full-duplex MAC protocol design and analysis," *IEEE Communications Letters*, vol. 19, no. 7, pp. 1185–1188, Jul. 2015.

[148] Y. Liao, B. Di, K. Bian, L. Song, D. Niyato, and Z. Han, "Cross-layer protocol design for distributed full-duplex network," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2015.

[149] K. Lin, R. H. Messerian, and Y. Wang, "A digital leakage cancellation scheme for monostatic FMCW radar," in *Proc. IEEE MTT-S International Microwave Symposium (IMS)*, Jun. 2004, pp. 747–750.

[150] K. Lin, Y. Wang, C. K. Pao, and Y. C. Shih, "A Ka-band FMCW radar front-end with adaptive leakage cancellation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 12, pp. 4041–4048, Dec. 2006.

[151] P. Lioliou and M. Viberg, "Least-squares based channel estimation for MIMO relays," in *Proc. International ITG Workshop on Smart Antennas (WSA)*, Feb. 2008, pp. 90–95.

[152] P. Lioliou, M. Viberg, and M. Coldrey, "Performance analysis of relay channel estimation," in *Proc. 43rd Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2009, pp. 1533–1537.

[153] P. Lioliou, M. Viberg, M. Coldrey, and F. Athley, "Self-interference suppression in full-duplex MIMO relays," in *Proc. 44th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2010, pp. 658–662.

[154] F. Liu, S. Jia, C. Guo, and C. Feng, "Exploiting polarization to resist phase noise for digital self-interference cancellation in full-duplex," in *Proc. IEEE International Conference on Communications (ICC)*, May 2016.

[155] G. Liu, F. R. Yu, H. Ji, V. C. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 500–524, Second Quarter 2015.

[156] Z. Luan, H. Qu, J. Zhao, and B. Chen, "Robust digital non-linear self-interference cancellation in full duplex radios with maximum correntropy criterion," *China Communications*, vol. 13, no. 9, pp. 53–59, Sep. 2016.

[157] D. G. C. Luck, *Frequency Modulated Radar*. McGraw-Hill, 1949.

[158] J. Ma, G. Y. Li, J. Zhang, T. Kuze, and H. Iura, "A new coupling channel estimator for cross-talk cancellation at wireless relay stations," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, Nov. 2009.

[159] D. P. Mandic and V. Su Lee Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*. Wiley, 2009.

[160] E. Manuzzato, J. Tamminen, M. Turunen, D. Korpi, F. Granelli, and M. Valkama, "Digitally-controlled electrical balance duplexer for transmitter-receiver isolation in full-duplex radio," in *Proc. 22nd European Wireless Conference (EW)*, May 2016, pp. 210–217.

[161] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, Mar. 2015.

[162] A. Masmoudi and T. Le-Ngoc, "A digital subspace-based self-interference cancellation in full-duplex MIMO transceivers," in *Proc. IEEE International Conference on Communications (ICC)*, Jun. 2015, pp. 4954–4959.

[163] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator for self-interference cancelation in full-duplex systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5122–5132, Jul. 2016.

[164] A. Masmoudi and T. Le-Ngoc, "Channel estimation and self-interference cancelation in full-duplex communication systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 321–334, Jan. 2017.

[165] "MAX2870 23.5 MHz to 6000 MHz fractional/integer-N synthesizer/VCO," Maxim Integrated, San Jose, California, USA.

[166] "SST12LP17A 2.4 GHz WLAN, high-efficiency power amplifier module," Microchip Technology Incorporated, Chandler, Arizona, USA.

[167] K. Min, S. Park, Y. Jang, T. Kim, and S. Choi, "Antenna ratio for sum-rate maximization in full-duplex large-array base station with half-duplex multiantenna users," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 10 168–10 173, Dec. 2016.

[168] "ZVE-8G+ coaxial amplifier," Mini-Circuits, Brooklyn, New York, USA.

[169] S. Mirabbasi and K. Martin, "Classical and modern receiver architectures," *IEEE Communications Magazine*, vol. 38, no. 11, pp. 132–139, Nov. 2000.

[170] M. Mohammadi, B. K. Chalise, H. A. Suraweera, C. Zhong, G. Zheng, and I. Krikidis, "Throughput analysis and optimization of wireless-powered multiple antenna full-duplex relay systems," *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1769–1785, Apr. 2016.

[171] M. Mohammadi, H. A. Suraweera, Y. Cao, I. Krikidis, and C. Tellambura, "Full-duplex radio for uplink/downlink wireless access with spatially random nodes," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 5250–5266, Dec. 2015.

[172] M. Mohammadi, H. A. Suraweera, and C. Tellambura, "Uplink and downlink rate analysis of a full-duplex C-RAN with radio remote head association," in *Proc. 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 778–782.

[173] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[174] M. Murad and A. M. Eltawil, "A simple full-duplex MAC protocol exploiting asymmetric traffic loads in WiFi systems," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2017.

[175] K. M. Nasr, J. P. Cosmas, M. Bard, and J. Gledhill, "Performance of an echo canceller and channel estimator for on-channel repeaters in DVB-T/H networks," *IEEE Transactions on Broadcasting*, vol. 53, no. 3, pp. 609–618, Sep. 2007.

[176] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.

[177] Nokia Solutions and Networks Oy, "5G masterplan – five keys to create the new communications era," white paper, 2016. [Online]. Available: https://resources.ext.nokia.com/asset/200316

[178] F. O'Hara and G. Moore, "A high performance CW receiver using feedthru nulling," *Microwave Journal*, vol. 6, pp. 63–71, Sep. 1963.

[179] Y. Pan, C. Zhou, G. Cui, W. Wang, and X. Li, "Self-interference cancellation with RF impairments suppression for full-duplex systems," in *Proc. 82nd IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2015.

[180] B. Paul, A. R. Chiriyath, and D. W. Bliss, "Survey of RF communications and sensing convergence research," *IEEE Access*, vol. 5, pp. 252–270, 2017.

[181] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Transactions on Communications*, vol. 55, no. 8, pp. 1607–1616, Aug. 2007.

[182] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[183] R.-A. Pitaval, O. Tirkkonen, R. Wichman, K. Pajukoski, E. Lähetkangas, and E. Tiirola, "Full-duplex self-backhauling for small-cell 5G networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 83–89, Oct. 2015.

[184] D. Pozar, *Microwave Engineering*. Wiley, 2012.

[185] C. Psomas, M. Mohammadi, I. Krikidis, and H. A. Suraweera, "Impact of directionality on interference mitigation in full-duplex cellular networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 487–502, Jan. 2017.

[186] P. Pursula, M. Kiviranta, and H. Seppä, "UHF RFID reader with reflected power canceller," *IEEE Microwave and Wireless Components Letters*, vol. 19, no. 1, pp. 48–50, Jan. 2009.

[187] X. Quan, Y. Liu, S. Shao, C. Huang, and Y. Tang, "Impacts of phase noise on digital self-interference cancellation in full-duplex communications," *IEEE Transactions on Signal Processing*, vol. 65, no. 7, pp. 1881–1893, Apr. 2017.

[188] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, V. Balan, and G. DeJean, "Rethinking indoor wireless mesh design: Low power, low frequency, full-duplex," in *Proc. Fifth IEEE Workshop on Wireless Mesh Networks (WiMesh)*, Jun. 2010.

[189] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Communications Surveys Tutorials*, vol. 15, no. 4, pp. 1567–1592, Fourth Quarter 2013.

[190] B. Razavi, "Design considerations for direct-conversion receivers," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, no. 6, pp. 428–435, Jun. 1997.

[191] B. Razavi, *RF microelectronics*. Prentice Hall, 1998.

[192] T. Riihonen, A. Balakrishnan, K. Haneda, S. Wyne, S. Werner, and R. Wichman, "Optimal eigenbeamforming for suppressing self-interference in full-duplex MIMO relays," in *Proc. 45th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2011.

[193] T. Riihonen, D. Korpi, O. Rantula, H. Rantanen, T. Saarelainen, and M. Valkama, "Inband full-duplex radio transceivers: A paradigm shift in tactical communications and electronic warfare?" *IEEE Communications Magazine*, vol. 55, no. 10, Oct. 2017.

[194] T. Riihonen, D. Korpi, O. Rantula, and M. Valkama, "On the prospects of full-duplex military radios," in *Proc. International Conference on Military Communications and Information Systems (ICMCIS)*, May 2017.

[195] T. Riihonen, P. Mathecken, and R. Wichman, "Effect of oscillator phase noise and processing delay in full-duplex OFDM repeaters," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2012, pp. 1947–1951.

[196] T. Riihonen, S. Werner, and R. Wichman, "Optimized gain control for single-frequency relaying with loop interference," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 2801–2806, Jun. 2009.

[197] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.

[198] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[199] T. Riihonen, S. Werner, R. Wichman, and E. Zacarias B., "On the feasibility of full-duplex relaying in the presence of loop interference," in *Proc. 10th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2009, pp. 275–279.

[200] T. Riihonen, R. Wichman, and J. Hämäläinen, "Co-phasing full-duplex relay link with non-ideal feedback information," in *Proc. IEEE International Symposium on Wireless Communication Systems (ISWCS)*, Oct. 2008, pp. 263–267.

[201] I. Rodriguez, H. C. Nguyen, N. T. K. Jørgensen, T. B. Sørensen, J. Elling, M. B. Gentsch, and P. Mogensen, "Path loss validation for urban micro cell scenarios at 3.5 GHz compared to 1.9 GHz," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013, pp. 3942–3947.

[202] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[203] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[204] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[205] A. Sahai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of full-duplex uplink/downlink channel," in *Proc. IEEE Information Theory Workshop (ITW)*, Sep. 2013.

[206] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "Understanding the impact of phase noise on active cancellation in wireless full-duplex," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2012, pp. 29–33.

[207] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4494–4510, Nov. 2013.

[208] H. Sakai, T. Oka, and K. Hayashi, "A simple adaptive filter method for cancellation of coupling wave in OFDM signals at SFN relay station," in *Proc. 14th European Signal Processing Conference (EUSIPCO)*, Sep. 2006.

[209] M. Sakai, H. Lin, and K. Yamashita, "Adaptive cancellation of self-interference in full-duplex wireless with transmitter IQ imbalance," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2014, pp. 3220–3224.

[210] K. Salehian, M. Guillet, B. Caron, and A. Kennedy, "On-channel repeater for digital television broadcasting service," *IEEE Transactions on Broadcasting*, vol. 48, no. 2, pp. 97–102, Jun. 2002.

[211] J. Sangiamwong, T. Asai, J. Hagiwara, Y. Okumura, and T. Ohya, "Joint multi-filter design for full-duplex MU-MIMO relaying," in *Proc. 69th IEEE Vehicular Technology Conference (VTC Spring)*, Apr. 2009.

[212] M. G. Sarret, M. Fleischer, G. Berardinelli, N. H. Mahmood, P. Mogensen, and H. Heinz, "On the potential of full duplex performance in 5G ultra-dense small cell networks," in *Proc. 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 764–768.

[213] M. Schoukens, R. Pintelon, and Y. Rolain, "Parametric identification of parallel Hammerstein systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 12, pp. 3931–3938, Dec. 2011.

[214] C. Y. A. Shang, P. J. Smith, G. K. Woodward, and H. A. Suraweera, "Linear transceivers for full duplex MIMO relays," in *Proc. Australian Communications Theory Workshop (AusCTW)*, Feb. 2014, pp. 11–16.

[215] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, 1948.

[216] S. Shao, X. Quan, Y. Shen, and Y. Tang, "Effect of phase noise on digital self-interference cancellation in wireless full duplex," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 2759–2763.

[217] A. Sharma, R. K. Ganti, and J. K. Milleth, "Joint backhaul-access analysis of full duplex self-backhauling heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1727–1740, Mar. 2017.

[218] P. F. Sielman, "A single-frequency communications system," *IEEE Transactions on Vehicular Technology*, vol. 23, no. 1, pp. 1–8, Feb. 1974.

[219] M. S. Sim, M. Chung, D. K. Kim, and C. B. Chae, "Low-complexity nonlinear self-interference cancellation for full-duplex radios," in *Proc. IEEE Globecom Workshops*, Dec. 2016.

[220] S. Simoens, O. Munoz-Medina, J. Vidal, and A. del Coso, "On the Gaussian MIMO relay channel with full channel state information," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3588–3599, Sep. 2009.

[221] R. Simpson and G. Tyler, "Radar scattering laws for the lunar surface," *IEEE Transactions on Antennas and Propagation*, vol. 30, no. 3, pp. 438–449, May 1982.

[222] N. Singh, D. Gunawardena, A. Proutiere, B. Radunovic, H. V. Balan, and P. Key, "Efficient and fair MAC for wireless networks with self-interference cancellation," in *Proc. International Symposium of Modeling and Optimization of Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, May 2011, pp. 94–101.

[223] "SKY65387-11 2110-2170 MHz variable gain amplifier," Skyworks Solutions Incorporated, Woburn, Massachusetts, USA.

[224] W. T. Slingsby and J. P. McGeehan, "A high-gain cell enhancer," in *Proc. 42nd IEEE Vehicular Technology Society Conference (VTS)*, May 1992, pp. 756–758.

[225] W. T. Slingsby and J. P. McGeehan, "Antenna isolation measurements for on-frequency radio repeaters," in *Proc. Ninth International Conference on Antennas and Propagation (ICAP)*, Apr. 1995, pp. 239–243.

[226] L. Song, Y. Liao, K. Bian, L. Song, and Z. Han, "Cross-layer protocol design for CSMA/CD in full-duplex WiFi networks," *IEEE Communications Letters*, vol. 20, no. 4, pp. 792–795, Apr. 2016.

[227] M. Steer, *Microwave and RF Design: A Systems Approach*. Scitech, 2010.

[228] A. G. Stove, "Linear FMCW radar techniques," *IEE Proceedings – Radar, Sonar and Navigation*, vol. 139, no. 5, pp. 343–350, Oct. 1992.

[229] Y. Sun, D. W. K. Ng, J. Zhu, and R. Schober, "Multi-objective optimization for robust power efficient and secure full-duplex wireless communication systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5511–5526, Aug. 2016.

[230] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 913–927, Feb. 2014.

[231] H. Suzuki, K. Itoh, Y. Ebine, and M. Sato, "A booster configuration with adaptive reduction of transmitter-receiver antenna coupling for pager systems," in *Proc. 50th IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 1999, pp. 1516–1520 vol.3.

[232] V. Syrjälä, M. Valkama, M. Allén, and K. Yamamoto, "Simultaneous transmission and spectrum sensing in OFDM systems using full-duplex radios," in *Proc. 82nd IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2015.

[233] V. Syrjälä, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 2977–2990, Jun. 2014.

[234] V. Syrjälä and K. Yamamoto, "Self-interference cancellation in full-duplex radio transceivers with oscillator phase noise," in *Proc. 20th European Wireless Conference (EW)*, May 2014, pp. 832–837.

[235] V. Syrjälä, "Analysis and mitigation of oscillator impairments in modern receiver architectures," Ph.D. dissertation, Tampere University of Technology, 2012.

[236] V. Syrjälä and M. Valkama, "Coexistence of LTE and WLAN in unlicensed bands: Full-duplex spectrum sensing," in *Proc. 10th International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM)*, 2015, pp. 725–734.

[237] H. Tabassum, A. H. Sakr, and E. Hossain, "Analysis of massive MIMO-enabled downlink wireless backhauling for full-duplex small cells," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2354–2369, Jun. 2016.

[238] J. Tamminen, M. Turunen, D. Korpi, T. Huusari, Y.-S. Choi, S. Talwar, and M. Valkama, "Digitally-controlled RF self-interference canceller for full-duplex radios," in *Proc. 24th European Signal Processing Conference (EUSIPCO)*, Aug. 2016, pp. 783–787.

[239] L. T. Tan, L. Ying, and D. W. Bliss, "Power control and relay selection in full-duplex cognitive relay networks: Coherent versus non-coherent scenarios," in *Proc. 51st Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2017.

[240] V. Tapio, M. Juntti, A. Pärssinen, and K. Rikkinen, "Real time adaptive RF and digital self-interference cancellation for full-duplex transceivers," in *Proc. 50th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2016, pp. 1558–1562.

[241] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 6, pp. 1510–1520, Jun. 2010.

[242] "CC2595 RF front-end transmit power amplifier for 2.4GHz ISM band systems," Texas Instruments Incorporated, Dallas, Texas, USA.

[243] L. Tian, S. Wang, Z. Cheng, and X. Bu, "All-digital self-interference cancellation in zero-IF full-duplex transceivers," *China Communications*, vol. 13, no. 11, pp. 27–34, Nov. 2016.

[244] L. Tomba, "On the effect of Wiener phase noise in OFDM systems," *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 580–583, May 1998.

[245] M. Valkama, M. Renfors, and V. Koivunen, "Advanced methods for I/Q imbalance compensation in communication receivers," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2335–2344, Oct. 2001.

[246] I. A. W. Vance, "Transmitter/receiver for single channel duplex communication system," U.S. Patent 4 238 850, Dec. 9, 1980.

[247] I. A. W. Vance, "Single channel duplex communication system," U.S. Patent 4 480 327, Oct. 30, 1984.

[248] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Transactions on Information Theory*, vol. 32, no. 1, pp. 85–96, Jan. 1986.

[249] B. Wang, J. Zhang, and A. Host-Madsen, "On the capacity of MIMO relay channels," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.

[250] D. Wang, R. Zhang, X. Cheng, L. Yang, and C. Chen, "Relay selection in full-duplex energy-harvesting two-way relay networks," *IEEE Transactions on Green Communications and Networking*, vol. 1, no. 2, pp. 182–191, Jun. 2017.

[251] L. Wang, K. Wu, and M. Hamdi, "Combating hidden and exposed terminal problems in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 11, no. 11, pp. 4204–4213, Nov. 2012.

[252] S. Wang, V. Venkateswaran, and X. Zhang, "Fundamental analysis of full-duplex gains in wireless networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 3, pp. 1401–1416, Jun. 2017.

[253] X. Wang, H. Huang, and T. Hwang, "On the capacity gain from full duplex communications in a large scale wireless network," *IEEE Transactions on Mobile Computing*, vol. 15, no. 9, pp. 2290–2303, Sep. 2016.

[254] D. Wen and G. Yu, "Time-division cellular networks with full-duplex base stations," *IEEE Communications Letters*, vol. 20, no. 2, pp. 392–395, Feb. 2016.

[255] D. Wen, G. Yu, R. Li, Y. Chen, and G. Y. Li, "Results on energy- and spectral-efficiency tradeoff in cellular networks with full-duplex enabled base stations," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1494–1507, Mar. 2017.

[256] D. Wu, C. Zhang, S. Gao, and D. Chen, "A digital self-interference cancellation method for practical full-duplex radio," in *Proc. IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Aug. 2014, pp. 74–79.

[257] X. Wu, Y. Shen, and Y. Tang, "The power delay profile of the single-antenna full-duplex self-interference channel in indoor environments at 2.6 GHz," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 1561–1564, Aug. 2014.

[258] H. Yang and T. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172–179, Feb. 2013.

[259] K. Yang, H. Cui, L. Song, and Y. Li, "Efficient full-duplex relaying with joint antenna-relay selection and self-interference suppression," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3991–4005, Jul. 2015.

[260] Y. Yang, B. Chen, K. Srinivasan, and N. B. Shroff, "Characterizing the achievable throughput in wireless networks with two active RF chains," in *Proc. IEEE Conference on Computer Communications (INFOCOM)*, Apr. 2014, pp. 262–270.

[261] A. Zeger, "Adaptive antenna same frequency repeater," in *Proc. Antennas and Propagation Society International Symposium (APSURSI)*, vol. 14, Oct. 1976, pp. 452–455.

[262] Y. Zeng and R. Zhang, "Full-duplex wireless-powered relay with self-energy recycling," *IEEE Wireless Communications Letters*, vol. 4, no. 2, pp. 201–204, Apr. 2015.

[263] C. Zhang, L. Laughlin, M. A. Beach, K. A. Morris, and J. L. Haine, "A self-interference cancellation testbed for full-duplex transceiver prototyping," in *Proc. 27th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2016.

[264] Y. P. Zhang and D. Liu, "Antenna-on-chip and antenna-in-package solutions to highly integrated millimeter-wave devices for wireless communications," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 10, pp. 2830–2841, Oct. 2009.

[265] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: Self-interference cancellation, protocol design, and relay selection," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 128–137, May 2015.

[266] Z. Zhang, K. Long, A. V. Vasilakos, and L. Hanzo, "Full-duplex wireless communications: Challenges, solutions, and future research directions," *Proceedings of the IEEE*, vol. 104, no. 7, pp. 1369–1409, Jul. 2016.

[267] Z. Zhang, Y. Shen, S. Shao, W. Pan, and Y. Tang, "Full duplex 2x2 MIMO radios," in *Proc. Sixth International Conference on Wireless Communications and Signal Processing (WCSP)*, Oct. 2014.

[268] Z. Zhang, X. Wang, K. Long, A. Vasilakos, and L. Hanzo, "Large-scale MIMO-based wireless backhaul in 5G networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 58–66, Oct. 2015.

[269] W. Zhao, C. Feng, F. Liu, C. Guo, and Y. Nie, "Polarization mismatch based self-interference cancellation against power amplifier nonlinear distortion in full duplex systems," in *Proc. 26th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Aug. 2015, pp. 256–260.

[270] C. Zhong, H. A. Suraweera, G. Zheng, I. Krikidis, and Z. Zhang, "Wireless information and power transfer with full duplex relaying," *IEEE Transactions on Communications*, vol. 62, no. 10, pp. 3447–3461, Oct. 2014.

[271] W. Zhou, K. Srinivasan, and P. Sinha, "RCTC: Rapid concurrent transmission coordination in full duplex wireless networks," in *Proc. 21st IEEE International Conference on Network Protocols (ICNP)*, Oct. 2013.

[272] F. Zhu, F. Gao, M. Yao, and H. Zou, "Joint information- and jamming-beamforming for physical layer security with full duplex base station," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6391–6401, Dec. 2014.

[273] Q. Zou, A. Tarighat, and A. H. Sayed, "Compensation of phase noise in OFDM wireless systems," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5407–5424, Nov. 2007.

Publications

Publication 1

D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014. DOI: 10.1109/TWC.2014.2315213

Full-Duplex Transceiver System Calculations: Analysis of ADC and Linearity Challenges

Dani Korpi, Taneli Riihonen, *Member, IEEE,* Ville Syrjälä, *Member, IEEE,* Lauri Anttila, *Member, IEEE,* Mikko Valkama, *Member, IEEE,* and Risto Wichman

***Abstract*—Despite the intensive recent research on wireless single-channel full-duplex communications, relatively little is known about the transceiver chain nonidealities of full-duplex devices. In this paper, the effect of nonlinear distortion occurring in the transmitter power amplifier (PA) and the receiver chain is analyzed, alongside with the dynamic range requirements of analog-to-digital converters (ADCs). This is done with detailed system calculations, which combine the properties of the individual electronics components to jointly model the complete transceiver chain, including self-interference cancellation. They also quantify the decrease in the dynamic range for the signal of interest caused by self-interference at the analog-to-digital interface. Using these system calculations, we provide comprehensive numerical results for typical transceiver parameters. The analytical results are also confirmed with full waveform simulations. We observe that the nonlinear distortion produced by the transmitter PA is a significant issue in a full-duplex transceiver and, when using cheaper and less linear components, also the receiver chain nonlinearities become considerable. It is also shown that, with digitally-intensive self-interference cancellation, the quantization noise of the ADCs is another significant problem.**

***Index Terms*—Full-duplex, direct-conversion transceiver, system calculations, nonlinear distortion, IIP2, IIP3, quantization noise, self-interference cancellation**

I. Introduction

FULL-DUPLEX (FD) radio technology, where the devices transmit and receive signals simultaneously at the same center-frequency, is the new breakthrough in wireless communications. Such frequency-reuse strategy can theoretically double the spectral efficiency, compared to traditional half-duplex (HD) systems, namely time-division duplexing (TDD) and frequency-division duplexing (FDD). Furthermore, since the transmission and reception happen at the same time at the same frequency, the transceivers can sense each other's transmissions and react to them. This, with appropriate medium access control (MAC) design, can result in a low level of signaling and low latency in the networks. Because of these benefits, full-duplex radios can revolutionize the design of radio communications networks.

D. Korpi, V. Syrjälä, L. Anttila, and M. Valkama are with the Department of Electronics and Communications Engineering, Tampere University of Technology, PO Box 692, FI-33101, Tampere, Finland, e-mail: dani.korpi@tut.fi.

T. Riihonen and R. Wichman are with the Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, PO Box 13000, FI-00076, Aalto, Finland.

The research work leading to these results was funded by the Academy of Finland (under the projects #259915, #258364 "In-band Full-Duplex MIMO Transmission: A Breakthrough to High-Speed Low-Latency Mobile Networks"), the Finnish Funding Agency for Technology and Innovation (Tekes, under the project "Full-Duplex Cognitive Radio"), the Linz Center of Mechatronics (LCM) in the framework of the Austrian COMET-K2 programme, and Emil Aaltonen Foundation.

However, there are still several problems in the practical realization and implementation of small and low-cost full-duplex transceivers. The biggest challenge is the so called self-interference (SI), which results from the fact that the transmitter and receiver use either the same [1], [2] or separate but closely-spaced antennas [3]–[6] and, thus, the transmit signal couples strongly to the receiver path. The power of the coupled signal can be, depending on, e.g., the antenna separation and transmit power, in the order of 60–100 dB stronger than the received signal of interest, especially when operating close to the sensitivity level of the receiver chain. In principle, the SI waveform can be perfectly regenerated at the receiver since the transmit data is known inside the device. Thus, again in principle, SI can be perfectly cancelled in the receiver path. However, because the SI signal propagates through an unknown coupling channel linking the transmitter (TX) and receiver (RX) paths, and is also affected by unknown nonlinear effects of the transceiver components, having perfect knowledge of the SI signal is, in practice, far from realistic.

In literature, some promising full-duplex radio demonstrations have recently been reported, e.g., in [3]–[6]. In these papers, both radio frequency (RF) and digital signal processing (DSP) techniques were proposed for SI suppression. Nearly 70 to 80 dB of overall attenuation has been reported at best, but in real-world scenarios SI mitigation results have not been even nearly that efficient [4]. This is because with low-cost small-size electronics, feasible for mass-market products, the RF components are subject to many nonidealities compared to idealized demonstration setups reported in [3]–[6].

Several recent studies have analyzed selected analog/RF circuit non-idealities in the context of practical full-duplex radios. The phase noise of the transmitter and receiver oscillators has been analyzed, e.g., in [7]–[10]. In these studies it was observed that the phase noise can potentially limit the amount of achievable SI suppression, especially when using two separate oscillators for transmitter and receiver. The effect of phase noise is also taken into account in the analysis presented in [11], where the feasibility of asynchronous full-duplex communications is studied. Furthermore, the impact of IQ mismatch induced mirror imaging has recently been addressd in [12].

In addition, the amplifiers and mixers cause nonlinear distortion, especially with transmit powers in the order of 10–50 dBm that are typical for, e.g., mobile cellular radios. This can have a big impact on the characteristics and efficient cancellation of the SI waveform. Nonlinear distortion is a particularly important problem in full-duplex radios, since the receiver RF components must be able to tolerate the high-power SI signal, which is then gradually suppressed in the

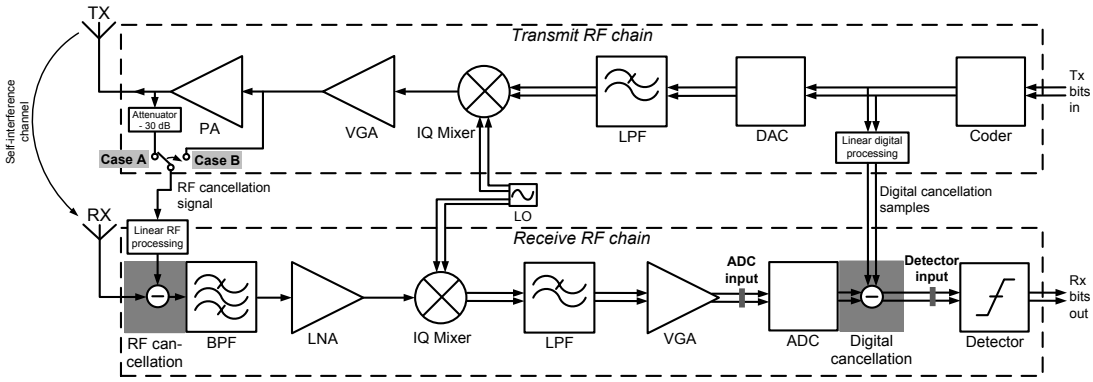


Fig. 1: Block diagram of the analyzed direct-conversion full-duplex transceiver, where RF and digital baseband interfaces for self-interference cancellation are illustrated in grey. Cases 'A' and 'B' refer to two alternative configurations assumed for reference signal extraction in RF cancellation.

RX chain. Recently, the effect of nonlinear distortion in a full-duplex transceiver, and its compensation, have been studied, e.g., in [13]–[16]. These studies indicate that nonlinear distortion of transceiver components, in particular with low-cost mass-product integrated circuits, forms a significant bottleneck in practical full-duplex radio devices.

Thus, in this paper, a comprehensive analysis of the nonlinear distortion effects in full-duplex transceivers is provided, with special focus on realistic achievable SI cancellation at receiver RF and DSP stages and corresponding maximum allowed transmit power. *Such analysis and understanding is currently missing from the literature of the full-duplex field.* The analysis covers the effects of both transmitter and receiver nonlinearities, and shows that both can easily limit the maximum allowed transmit power of the device. Explicit expressions are provided that quantify the overall second- and third-order nonlinear distortion power, due to all essential RF components, at the detector input in the receiver. These can be used directly to, e.g., derive the required linearity figures for the transceiver RF components such that the nonlinear distortion at detector input is within any given implementation margin.

We also analyze, quantify, and compare two alternative RF cancellation strategies where reference signal is taken either from TX power amplifier (PA) input or output. We then show that PA nonlinearity can seriously limit the device operation already with transmit powers in the order of 5–10 dBm, especially when RF cancellation reference is taken from PA input. This indicates that, in addition to RX path, the linearity of the TX chain is also of high concern when designing and implementing full-duplex transceivers. The effect of transmit imperfections is also analyzed in [17]–[20] with a relatively simplified model. However, in this paper, the analysis of the transmit imperfections is done based on the actual properties of the TX components.

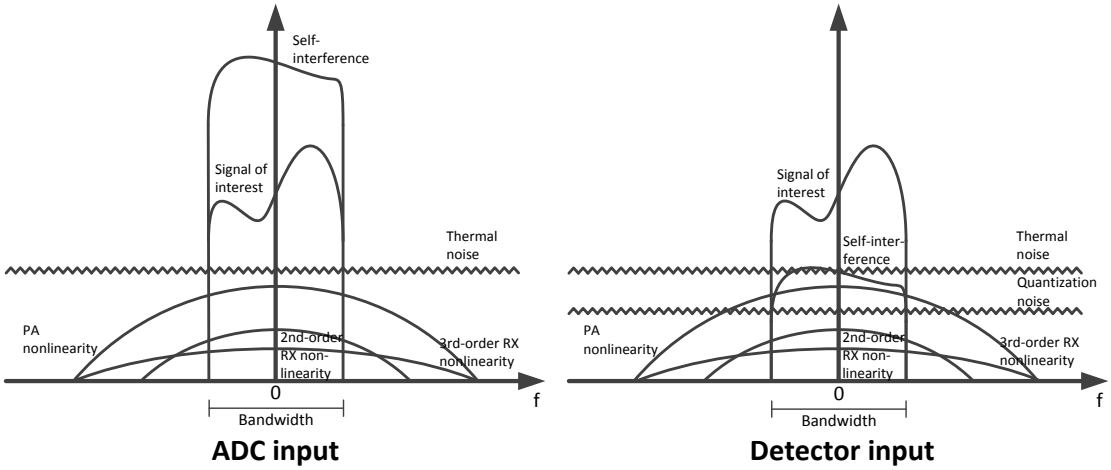
Finally, in addition to linearity analysis, the required dynamic range of the analog-to-digital converter (ADC) is addressed in this paper. Since a considerable amount of the SI cancellation is carried out in the digital domain, additional dynamic range is needed in the analog-to-digital interface or otherwise the SI signal heavily decreases the effective resolution of the weak desired signal. This, in turn, limits the performance of the whole transceiver. In this paper, we will explicitly quantify and derive the ADC dynamic range and resolution requirements such that the signal-to-interference-plus-noise ratio (SINR) at detector input in the RX chain does not degrade more than the specified implementation margin. *Such analysis is also missing from the literature*. In particular, earlier work in [21] focuses on ADCs within an otherwise ideal system, while the current analysis incorporates the joint effect of quantization noise and all other nonidealities.

The organization of the rest of the paper is as follows. Section II describes the analyzed full-duplex direct-conversion transceiver model and especially the nonlinear characteristics of the essential TX and RX components. The system calculations, in terms of the powers of the useful and interfering signal components in different stages of the RX chain, as well as the required ADC performance, are then carried out and analyzed in Section III. Section IV provides the actual waveform-level reference simulation results of a complete full-duplex device, verifying the good accuracy of the system calculations and the associated performance limits. Finally, conclusions are drawn in Section V.

Nomenclature: Throughout the paper, the use of linear power units is indicated by lowercase letters. Correspondingly, when referring to logarithmic power units, uppercase letters are used. The only exception to this is the noise factor, which is denoted by capital F according to common convention in the literature of the field. Watts are used as the absolute power unit, and dBm as the logarithmic power unit.

II. Full-duplex Transceiver Modeling

Our approach is to model a complete full-duplex transceiver component-wise, which allows us to analyze the feasibility of single-frequency full-duplex communications. Most of the emphasis in the calculations is at the receiver side since, due



(a) A sketch of the signal spectra at the ADC input.

(b) A sketch of the signal spectra at the detector input.

Fig. 2: A principal illustration of the signal spectra at the inputs of the ADC and the detector. Note that this figure depicts a situation with a medium-level transmit power. With higher transmit powers, nonlinear distortion is more likely to be the dominant distortion component.

to the powerful self-interference, it is the more delicate part of the transceiver in terms of enabling full-duplex operation. Nevertheless, the effects of the transmitter are also taken into account as the exact SI waveform depends on, e.g., power amplifier nonlinearities. We wish to again emphasize that also oscillator phase noise can represent a performance bound in FD devices [7]–[10]. However, as the focus in this article is on nonlinear distortion and ADC interface, phase noise is neglected in the following.

A block diagram representing the analyzed full-duplex direct-conversion transceiver is given in Fig. 1. For generality, both RF- and DSP-based SI cancellation [21] are covered in the analysis. The direct-conversion architecture is chosen due to its simple structure and wide applications, e.g., in cellular devices. Another signifcant aspect is the assumed reference signal path for RF cancellation. In this paper, two alternative scenarios are analyzed: Case A, in which the reference signal is taken from the output of the PA and attenuated to a proper level, and Case B, in which the reference signal is taken directly from the input of the PA. These scenarios are also marked in the block diagram in Fig. 1 using a switch. In general, both RF and DSP cancellation stages are assumed to deploy only linear processing.

A. Analysis Principles and Performance Measures

In the transceiver system calculations, the two most relevant interfaces are the *ADC input* and *detector input*. These points are also marked in the block diagram in Fig. 1. Furthermore, example signal characteristics and the different signal components, alongside with their typical relative power levels, are illustrated in Figs. 2a and 2b. The reason for the significance of the ADC input is the role of quantization and its dependence on SI. As the receiver automatic gain control (AGC) keeps the total ADC input at constant level, higher SI power means reduced desired signal power and thus more and more of the ADC dynamic range is reserved by the SI signal. This, in turn, indicates reduced effective resolution for the desired signal, which may limit the receiver performance.

The effect of quantization is studied by determining the SINR at the ADC input, quantifying the power of the desired signal relative to the other signal and distortion components at this point. A typical situation in terms of the power levels at this interface can be seen in Fig. 2a, where the SI signal is clearly dominating, and thus reserving a signicant amount of dynamic range.

Then, to characterize the overall performance of the whole full-duplex transceiver, and how the different types of distortion affect it, also the final SINR at the detector input, including digital SI cancellation, is studied and analyzed. This is thus the other significant point or calculation interface in the forthcoming analysis. Typical power levels also at this interface can be seen in Fig. 2b, where the SI signal has now been attenuated by digital cancellation, and it is no more the dominant distortion component. However, due to analog-to-digital conversion, there is now quantization noise in the total signal, which might be a significant issue, depending on the parameters of the transceiver.

Throughout the rest of the article, it is assumed that all the distortion types can be modelled in additive form. This is very typical in transceiver system calculations, see, e.g., [22], [23]. The good accuracy of this approach is also verified by full waveform simulations later in Section IV.

Under the above assumptions, the SINR on linear scale at the ADC input can now be directly defined as

$$sinr_{\mathrm{ADC}} = \frac{g_{\mathrm{rx}} p_{\mathrm{SOI,in}}}{g_{\mathrm{rx}} F p_{\mathrm{N,in}} + \frac{g_{\mathrm{rx}}}{a_{\mathrm{ant}}} \left(\frac{p_{\mathrm{tx}}}{a_{\mathrm{RF}}} + \frac{p_{\mathrm{3rd,PA,tx}}}{a_{\mathrm{NL}}} \right) + p_{\mathrm{2nd}} + p_{\mathrm{3rd}}}, \tag{1}$$

where g_{rx} is the total gain of the RX chain, $p_{\mathrm{SOI,in}}$ is the power

of the signal of interest at RX input, F is the noise factor of the receiver, $p_{\text{N,in}}$ is the thermal noise power at the input of the receiver, a_{ant} and a_{RF} are the amounts of antenna attenuation and RF cancellation, p_{tx} is the transmit power, $p_{\text{3rd,PA,tx}}$ is the power of PA-induced nonlinear distortion at the output of the transmit chain, parameter a_{NL} is a_{RF} for Case A and 1 for Case B, and p_{2nd} and p_{3rd} are the cumulated powers of 2nd- and 3rd-order nonlinear distortion produced at the RX chain. All the powers are assumed to be in linear units, which is indicated also by the lowercase symbols. These signal components are illustrated in Fig. 2a with realistic relative power levels.

The purpose of defining the ADC input SINR is to quantify the ratio of the useful signal power and total noise-plus-interference power entering the analog-to-digital interface. With fixed ADC voltage range, and assuming that the overall receiver gain is controlled properly, the total ADC input power $g_{\text{rx}}p_{\text{SOI,in}} + g_{\text{rx}}Fp_{\text{N,in}} + \frac{g_{\text{rx}}}{a_{\text{ant}}a_{\text{RF}}}p_{\text{tx}} + \frac{g_{\text{rx}}}{a_{\text{NL}}}p_{\text{3rd,PA,tx}} + p_{\text{2nd}} + p_{\text{3rd}}$ is always matched to the maximum allowed average power, say p_{target}. This will be elaborated in more details later.

Taking next the quantization noise and digital cancellation into account, the SINR at the detector input can be defined as

$$sinr_{\text{D}} = \frac{g_{\text{rx}}p_{\text{SOI,in}}}{g_{\text{rx}}Fp_{\text{N,in}} + \frac{g_{\text{rx}}}{a_{\text{ant}}}\left(\frac{p_{\text{tx}}}{a_{\text{RF}}a_{\text{dig}}} + \frac{p_{\text{3rd,PA,tx}}}{a_{\text{NL}}}\right) + p_{\text{quant}} + \cdots + p_{\text{2nd}} + p_{\text{3rd}}}, \tag{2}$$

where a_{dig} is the attenuation achieved by digital cancellation and p_{quant} is the power of quantization noise. This SINR defines the overall receiver performance of the full-duplex transceiver and is thus the most significant figure of merit in the analysis. A realistic scetch of the relative power levels of the specified signal components also at this interface can be seen in Fig. 2b.

The following subsections analyze in detail the different component powers of the above two principal equations, and their dependence on the transmit power, RF cancellation, digital cancellation, and TX and RX chain nonlinear characteristics. Then, in Section III, these are all brought together and it is analyzed in detail how these elementary parameters and transceiver characteristics affect the SINR at both of the studied interfaces and thereon the whole transceiver operation.

B. Radio-Frequency Front-End

1) Receiver Reference Sensitivity: The most challenging situation from the SI suppression perspective is when the actual received signal is close to the receiver sensitivity level. Thus, we begin by briefly defining the receiver reference sensitivity, which is determined by the thermal noise floor at RX input, the noise figure of the receiver, and the signal-to-noise ratio (SNR) requirement at the detector. This forms then the natural reference for assumed received signal levels in our analysis. The reference sensitivity, expressed in dBm, follows directly from [22] and can be written as

$$P_{\text{sens}} = -174 + 10\log_{10}(B) + NF_{\text{rx}} + SNR_{\text{d}}, \tag{3}$$

where B is the bandwidth of the system in Hertz, NF_{rx} is the noise figure of the receiver, and SNR_{d} is the SNR requirement at the input of the detector. In modern radio systems, the sensitivity is, strictly-speaking, affected by the assumed code rate and modulation through varying SNR requirements. However, for simplicity, only two reference sensitivity numbers are assumed in this study in the numerical examples of Section III.

The total receiver noise figure, in dB, is in general defined as $NF_{\text{rx}} = 10\log_{10}(F_{\text{rx}})$ where the total noise factor of the assumed RX chain in Fig. 1 is given by the classical Friis' formula [22] as

$$F_{rx} = F_{\text{LNA}} + \frac{F_{\text{mixer}} - 1}{g_{\text{LNA}}} + \frac{F_{\text{VGA}} - 1}{g_{\text{LNA}}g_{\text{mixer}}}. \tag{4}$$

In above, F_{LNA}, F_{mixer}, and F_{VGA} are the noise factors of the LNA, IQ Mixer, and VGA, respectively. Similarly, g_{LNA}, g_{mixer}, and g_{VGA} are the linear gains of the components.

2) RF Cancellation: In general, depending on the antenna separation, the path loss between the transmit and receive antennas attenuates the SI signal to a certain degree. However, to prevent the saturation of the RX chain, additional RF cancellation is most likely required. For generality, a multi-tap RF cancellation circuit, as presented in [13], [24], is assumed in this paper. This type of a cancellation circuit consists of several fixed delay lines, each of which has its own weight factor. Thus, the final cancellation signal consists of a linear combination of several delayed versions of the reference transmit signal with appropriate phase and amplitude tuning. The cancellation is then done by estimating the coefficients for the different delay lines based on the SI coupling channel, and subtracting this cancellation signal from the received signal. Thus, depending on the chosen delays, this type of a RF cancellation scheme might even be able to attenuate the multipath components. It should be noted, however, that in terms of the actual system calculations there is no difference between using a single- or multi-tap RF canceller, as only the amount of achieved SI attenuation is taken into account by the equations.

Furthermore, in our analysis, two alternatives for the reference signal path are considered, as follows.

- Case A describes perhaps the most widely used implementation technique for taking the reference signal for RF cancellation [3], [4], [25]–[27]. However, the drawback of this approach is the need for a bulky RF attenuator to achieve a feasible power level for the cancellation signal. The required amount of attenuation is obviously the estimated path loss between the antennas, as this ensures that the powers of the reference signal and incoming SI signal are of similar magnitude at the RF cancellation block.
- In Case B, the reference signal is taken already from the input of the PA. As the gain of the PA is usually within 10 dB of the magnitude of the path loss between the antennas [22], [23], only a tunable amplitude and phase matching circuit, such as a commercial product [28], with feasible tuning range is required. Thus, no additional RF attenuator is needed, resulting in a simpler and lower-cost RF frontend. On the other hand, as shown in this paper, the problem in this implementation is the nonlinear distortion produced by the PA, which is not included in

the reference signal. Thus, it is not attenuated by RF cancellation like in Case A, resulting in a lower SINR in the analog domain. This will be illustrated in Section III. Notice also that from the PA nonlinearity perspective, Case B is equivalent to the method used in [5] and [29], where a separate low-power TX chain is used to generate the RF reference signal. Thus, in our analysis, Case B covers indirectly also this type of transceiver scenarios.

C. Analog-to-Digital Interface and Digital Cancellation

Next we address issues related to analog-to-digital interface and quantization noise, especially from the perspective of residual SI left for digital cancellation. The starting point is the classical expression, available in, e.g., [22], defining the signal-to-quantization-noise ratio of the ADC as

$$SNR_{\text{ADC}} = 6.02b + 4.76 - PAPR, \tag{5}$$

where b is the number of bits at the ADC, and $PAPR$ is the estimated peak-to-average power ratio. The above expression assumes proper AGC at ADC input such that the full range of the ADC is used but the clipping of the signal peaks is avoided. However, the analysis could be easily translated to cover clipping noise as well [21].

Building on the above expression, our approach to analyze the impact of SI on analog-to-digital interface is to determine how many bits are effectively lost from the signal of interest. This is directly based on the fact that the remaining SI signal reserves part of the dynamic range of the ADC and thus decreases the resolution of the desired signal. Now, the amount of lost bits due to RX noise and interference can be determined by calculating how many dBs the signal of interest is below the total signal power, as this is directly the amount of dynamic range that is reserved by the noise and interference. The amount of lost bits can thus in general be calculated from (5) as

$$b_{\text{lost,I+N}} = \frac{P_{\text{tot}} - P_{\text{SOI}}}{6.02}, \tag{6}$$

where P_{tot} and P_{SOI} are the total power of the signal and the power of the signal of interest at the input of the ADC, respectively, and 6.02 depicts the dynamic range of one bit, thus mapping the loss of dynamic range to loss of bits. Then, the actual *bit loss due to self-interference* is defined as the increase in lost bits when comparing the receiver operation with and without SI. Following this step-by-step path, and using (6), a closed-form equation for the bit loss can be derived as shown in detail in Appendix A, yielding

$$b_{\text{lost}} = \log_4 \left[1 + \left(\frac{1}{p_{\text{SOI,in}} + p_{\text{N,in}}} \right) \cdot \left(\frac{p_{\text{tx}}}{a_{\text{ant}} a_{\text{RF}}} + \frac{p_{\text{tx}}^3}{a_{\text{ant}} a_{\text{NL}} iip3_{\text{PA}}^2 g_{\text{PA}}^2} \right) \right]. \tag{7}$$

Here, $iip3_{\text{PA}}$ and g_{PA} are the IIP3 figure and gain of the PA in linear units, respectively.

An immediate observation following from (7) is that increasing the transmit power with respect to the other signal components also increases the bit loss. Furthermore, increasing antenna attenuation or RF cancellation decreases the bit loss. These are relatively intuitive results, but with (7) they can be quantified and analyzed exactly. It is also important to note that the bit loss does not depend on the total amount of bits in the ADC. Thus, the detailed numerical illustrations given in Section III, based on (7), apply to all ADCs.

Finally, prior to detection, the remaining SI is mitigated in the digital domain by subtracting the transmitted baseband waveform from the received signal. The subtracted samples are generated by linearly filtering the transmitted symbols with an estimate of the overall coupling channel response linking the TX and RX. In practice, the channel estimation at this stage includes the effects of the transmitter, the coupling channel between the antennas, and the receiver. Also the multipath components due to reflections are included in the channel estimate. In our analysis, as was already illustrated in (2), the efficiency of digital cancellation is parameterized through digital SI attenuation a_{dig}, or A_{dig} in dB. Notice that since only linear digital cancellation is assumed, only the linear SI component is suppressed.

D. Nonlinear Distortion in Receiver Chain

In addition to quantization noise, the nonlinear distortion produced by the components of the transceiver is also of great interest. Following the well-established conventions from literature, nonlinear distortion of individual components is modeled by using the IIP2 and IIP3 figures (2nd- and 3rd-order input-referred intercept points) [22]. For a general nth-order nonlinearity, the power of the nonlinear distortion in dBm at the output of the component is given by

$$P_{\text{nth}} = P_{\text{out}} - (n-1)(IIPn - P_{\text{in}}), \tag{8}$$

where P_{in} is the total input power of the component, P_{out} is the total output power, and $IIPn$ is nth-order input-referred intercept point, all in dBm. As is well known in the literature, such principal power characteristics apply quite accurately, given that the component is not driven to full saturation, while offering analytically tractable expressions to accumulate total nonlinear distortion powers of a complete transceiver chain. In the case of the RX chain, this includes the LNA, mixers, and baseband VGA. The accuracy of this approach over a wide range of parameters, e.g., transmit powers, is illustrated and verified through full reference waveform simulations in Section IV.

E. Transmitter Modeling and PA Nonlinearity

When analyzing and modeling the TX chain, it is assumed that the power of thermal noise is negligibly low. This is a reasonable assumption as transmitters are never limited by inband thermal noise floor. Hence, thermal noise is omitted from transmitter modeling and only injected at RX input. Furthermore, we also assume that the power amplifier is the main source of nonlinear distortion, since all other transmitter components operate at low power regime. In fact, even if some nonlinear distortion was created, e.g., in the feeding amplifier prior to PA, it is a part of the RF cancellation reference signal in all the considered scenarios, and hence suppressed by RF

cancellation below the RX noise level. Thus, it is sufficient to focus on the nonlinearities of the PA when analyzing the transmitter.

The PA itself, in turn, is typically heavily nonlinear [22], [23], [30]. In our analysis, we assume that the PA produces 3rd-order distortion which falls on to the signal band, since this is the dominant distortion in practice. This is characterized with the IIP3 figure of the PA, according to (8). Furthermore, in Case A, this distortion is included in the reference signal, and is thus attenuated by RF cancellation. In Case B, this is not the case, and the nonlinear distortion produced by the PA remains at the same level after RF cancellation, as it is only attenuated by the coupling channel path loss.

Another observation about the nonlinearities of the transmit chain is that linear digital cancellation cannot suppress them, because the reference symbols for digital cancellation exist only in the digital domain and do not include any analog distortion. Moreover, nonlinear distortion cannot be modelled with a linear filter, and thus linear digital cancellation is unable to mitigate it. *The results shown in this paper thus give motivation to develop nonlinear digital SI cancellation techniques*. First works to this direction have been very recently reported in [13]–[15].

F. Accumulated Component Powers at Detector Input

The previous subsections describe elementary component-level modeling principles. Next, in this subsection, we accumulate the total observable power levels of all essential individual signal components at the *input of the detector*. This includes the desired signal power, (residual) SI power, quantization noise power, thermal noise power, RX 2nd- and 3rd-order nonlinear distortion power, and TX PA induced 3rd-order nonlinear distortion power.

First, the power of the quantization noise at the detector input can be written as

$$P_{\text{quant}} = P_{\text{target}} - SNR_{\text{ADC}} = P_{\text{target}} - 6.02b - 4.76 + PAPR, \tag{9}$$

where P_{target} is the maximum allowed average power of the signal at the ADC input, such that clipping is avoided. For any given PAPR, it can be observed that the power of the quantization noise depends only on the characteristics of the ADC, namely its maximum input power and the amount of bits.

The powers of the other signal components depend on several parameters, first and foremost on the total gain of the RX chain. As the signal of interest, SI signal, and the nonlinear distortion produced by the PA are the only significant signal components at the very input of the receiver, the total gain in linear units can be first written as

$$g_{\text{rx}} = \frac{p_{\text{target}}}{\frac{1}{a_{\text{ant}}}\left(\frac{p_{\text{tx}}}{a_{\text{RF}}} + \frac{p_{\text{3rd,PA,tx}}}{a_{\text{NL}}}\right) + p_{\text{SOI,in}}}. \tag{10}$$

When considering Case A, the nonlinear distortion produced by the PA is attenuated by RF cancellation. Thus, with high transmit powers, the power of the total signal at the input of the receiver can be approximated by the power of SI, as it is several orders of magnitude higher than the power of any other signal component when operating close to the sensitivity level. In this case, (10) is simplified to

$$g_{\text{rx}} = \frac{a_{\text{ant}} a_{\text{RF}} p_{\text{target}}}{p_{\text{tx}}}. \tag{11}$$

Knowing now the total gain of the receiver, it is then trivial to write the expressions for the powers of the other signal components, namely the signal of interest and thermal noise, at the input of the detector in dBm as

$$P_{\text{SOI}} = P_{\text{SOI,in}} + G_{\text{rx}} \text{ and} \tag{12}$$

$$P_{\text{N}} = P_{\text{N,in}} + G_{\text{rx}} + NF_{\text{rx}}. \tag{13}$$

The corresponding power of linear SI can be written as

$$P_{\text{SI}} = P_{\text{tx}} - A_{\text{ant}} - A_{\text{RF}} - A_{\text{dig}} + G_{\text{rx}}. \tag{14}$$

Furthermore, for high transmit powers, when (11) can be used to approximate the total gain of the RX chain, (14) becomes $P_{\text{SI}} = P_{\text{target}} - A_{\text{dig}}$.

Next, the total powers of the 2nd- and 3rd-order nonlinear distortion, produced by the RX chain, are derived based on (8), as shown in detail in Appendix B. The resulting equations are

$$p_{\text{2nd}} \approx g_{\text{LNA}}^2 g_{\text{mixer}} g_{\text{VGA}} p_{\text{in}}^2 \left(\frac{1}{iip2_{\text{mixer}}} + \frac{g_{\text{mixer}}}{iip2_{\text{VGA}}}\right) \tag{15}$$

$$p_{\text{3rd}} \approx g_{\text{LNA}} g_{\text{mixer}} g_{\text{VGA}} p_{\text{in}}^3 \left[\left(\frac{1}{iip3_{\text{LNA}}}\right)^2 + \left(\frac{g_{\text{LNA}}}{iip3_{\text{mixer}}}\right)^2 + \left(\frac{g_{\text{LNA}} g_{\text{mixer}}}{iip3_{\text{VGA}}}\right)^2\right], \tag{16}$$

where the subscript of each parameter indicates the considered component. Furthermore, $iip2_{\text{k}}$ and $iip3_{\text{k}}$ are the 2nd- and 3rd-order input intercept points expressed in Watts, g_{k} is the linear gain of the corresponding component, and p_{in} is the total power of the signal after RF cancellation, again in Watts.

Finally, the power of the PA-induced nonlinear distortion at the output of the transmit chain can be written as

$$\begin{aligned} P_{\text{3rd,PA,tx}} &= P_{\text{tx}} - 2(IIP3_{\text{PA}} - (P_{\text{tx}} - G_{\text{PA}})) \\ &= 3P_{\text{tx}} - 2(IIP3_{\text{PA}} + G_{\text{PA}}), \end{aligned} \tag{17}$$

This value is used in, for example, (10), as the gain is determined based on the signal levels at the input of the RX chain. The power of the PA-induced nonlinear distortion at the input of the detector can then be written as

$$\begin{aligned} P_{\text{3rd,PA}} &= P_{\text{3rd,PA,tx}} + G_{\text{rx}} - A_{\text{ant}} - A_{\text{NL}} \\ &= 3P_{\text{tx}} - 2(IIP3_{\text{PA}} + G_{\text{PA}}) + G_{\text{rx}} - A_{\text{ant}} - A_{\text{NL}}. \end{aligned} \tag{18}$$

As only linear digital cancellation is deployed, the nonlinear distortion produced by the PA is only attenuated by the coupling channel path loss (A_{ant}), and potentially by RF cancellation ($A_{\text{NL}} = A_{\text{RF}}$), if considering Case A. Further attenuation of this nonlinear component with actual nonlinear cancellation processing, analog or digital, is out of the scope of this paper. The potential benefits of digitally attenuating nonlinearly distorted SI signals are analyzed in, e.g., [13]–[15].

III. System Calculations and Results

In this section, we put together the elementary results of the previous section in terms of overall system calculations. The basic assumption is that the actual received signal is only slightly above the receiver sensitivity level, as this is the most challenging case from the SI perspective. The main interests of these calculations are then to see how much the quantization noise produced by the ADC affects the overall performance of the transceiver, and how severe the nonlinear distortion products, caused by full-duplex operation, are at the detector input. For this reason, the final signal quality ($SINR_{\mathrm{d}}$) after the ADC and digital cancellation is measured with different parameters and transmit powers.

In all the experiments, the maximum allowed SINR loss due to full-duplex operation is assumed to be 3 dB. This means that if the effective total noise-plus-interference power more than doubles compared to classical half-duplex operation, then the receiver performance loss becomes too high. Thus, the derived $SINR_{\mathrm{d}}$ values under FD operation are compared to signal-to-thermal-noise-ratio (SNR_{d}) at the input of the detector. The transmit power level at which this 3 dB loss is reached is referred to as the *maximum transmit power*. It is marked to all relevant result figures with a vertical line to illustrate what is effectively the highest transmit power with which the full-duplex transceiver can still operate with tolerable SINR loss.

This also provides a way to obtain some insight into the relative performances of half-duplex and full-duplex radio devices. Namely, with a low SINR loss, a full-duplex radio can be assumed to approximately double the spectral efficiency, whereas with a high SINR loss, the effective spectral efficiency might be even lower than that achieved by traditional half-duplex radios. A SINR loss of 3 dB illustrates a point at which full-duplex transceivers can still be expected to provide a capacity gain in comparison to half-duplex transceivers. In addition, (7) compares the effect of quantization noise in full-duplex and half-duplex transceivers by determining the SI-induced decrease in the effective dynamic range of the ADC. However, a more in-depth analysis regarding the performance of practical full-duplex transceivers, especially at system/network level, is out of the scope of this paper, and we consider it as a possible topic for future work. Furthermore, a detailed performance comparison between half-duplex and full-duplex radios under some implementation impairments, excluding nonlinear distortion, is already done in [31].

A. Parameters for Numerical Results

In order to provide actual numerical results with the derived equations, parameters for the full-duplex transceiver are specified. It should be emphasized that the chosen parameters are just example numbers chosen for illustration purposes only, and all the calculations can be easily repeated with any given parametrization.

1) Receiver: The general system level parameters of the studied full-duplex transceiver are shown in Table I, and the parameters of the individual components of the receiver are shown in Table II. Two sets of parameters are used, which are referred to as Parameter Set 1 and Parameter Set 2. The first set of parameters corresponds to state-of-the-art wideband RF transceiver performance. The parameters of the 2nd set model a more challenging scenario with lower received signal power, decreased linearity, and slightly inferior SI cancellation ability. In most parts of the analysis, Parameter Set 1 is used as it depicts better the characteristics of modern transceivers, especially in terms of bandwidth and linearity.

TABLE I: System level parameters of the full-duplex transceiver for Parameter Sets 1 and 2.

Parameter	Value for Param. Set 1	Value for Param. Set 2
SNR requirement	10 dB	5 dB
Bandwidth	12.5 MHz	3 MHz
Receiver noise figure	4.1 dB	4.1 dB
Sensitivity	-88.9 dBm	-100.1 dBm
Received signal power	-83.9 dBm	-95.1 dBm
Antenna separation	40 dB	40 dB
RF cancellation	40 dB	20 dB
Digital cancellation	35 dB	35 dB
ADC bits	8	12
ADC P-P voltage range	4.5 V	4.5 V
PAPR	10 dB	10 dB
Allowed SINR loss	3 dB	3 dB

TABLE II: Parameters for the components of the receiver. The values in the parentheses are the values used in Parameter Set 2.

Component	Gain [dB]	IIP2 [dBm]	IIP3 [dBm]	NF [dB]
BPF	0	-	-	0
LNA	25	43	-9 (-15)	4.1
Mixer	6	42	15	4
LPF	0	-	-	0
VGA	0–69	43	14 (10)	4
Total	31–100	11	-17 (-21)	4.1

With (3), the sensitivity level of the receiver can be calculated as $P_{\mathrm{sens}} = -88.9$ dBm for Parameter Set 1. This is a typical realistic value and close to the reference sensitivity specified in the LTE specifications [32]. For Parameter Set 2, the corresponding sensitivity is $P_{\mathrm{sens}} = -100.1$ dBm, which is an even more challenging value, assuming that the power of the received signal is close to the sensitivity level. Here, the power of the received signal is assumed to be 5 dB above sensitivity level, resulting in a received power level of either $P_{\mathrm{SOI,in}} = -83.9$ dBm or $P_{\mathrm{SOI,in}} = -95.1$ dBm, depending on the parameter set.

The isolation between the antennas is assumed to be 40 dB. This value, or other values of similar magnitude, have been reported several times in literature [4]–[6]. Furthermore, the assumed RF cancellation level for Parameter Set 1 is 40 dB. For a single-tap RF canceller (used, e.g., in [3], [4]), this value is somewhat optimistic However, if a multi-tap RF cancellation circuit is considered, RF cancellation values of this magnitude can easily be expected [13]. In Parameter Set 2, in turn, a lower RF cancellation level of 20 dB is assumed to represent a more practical scenario.

The component parameters of the actual direct-conversion RX chain are determined according to [33]–[35]. The objective

TABLE III: Parameters for the components of the transmitter.

Component	Gain [dB]	IIP3 [dBm]	NF [dB]
LPF	0	-	0
Mixer	5	5	9
VGA	0–35	5	10
PA	27	20	5
Total	32–67	−20	10.3

is to select typical parameters for each component, and thus obtain reliable and feasible results. The chosen parameters are shown in Table II, where the values without parentheses are used with Parameter Set 1, while the values with parentheses are used with Parameter Set 2. With (15) and (16), the total IIP2 and IIP3 figures of the whole receiver can be calculated to be 10.8 dBm and -17.1 dBm (Parameter Set 1) or 10.8 dBm and -20.1 dBm (Parameter Set 2), respectively.

The ADC input is controlled by the VGA such that the assumed full voltage range is properly utilized. As a realistic scenario, PAPR of the total signal is assumed to be 10 dB and state-of-the-art ADC specifications in [36] are deployed in terms of full voltage range. Using now (5), the signal-to-quantization noise ratio of the ADC is $SNR_{\text{ADC}} = 6.02b - 5.24$, where b is the number of bits at the ADC.

2) Transmitter: The parameters of the individual TX components are shown in Table III, and they are the same for both parameter sets. Again, typical values are chosen for the parameters according to [22] and [23]. This ensures that the conclusions apply to a realistic TX chain. Furthermore, for the transmitter, only 3rd-order nonlinear distortion is taken into account as the 2nd-order nonlinearities do not fall on the actual signal band. Assuming that the power of the feeding amplifier input signal is approximately -35 dBm, it can be observed from the table that, with the maximum feeding amplifier gain, the power of the 3rd-order nonlinear distortion at the output of the transmit chain is 40 dB lower than the fundamental signal component. Hence, the spectral purity of the considered TX chain is relatively high, and thus the obtained results, when it comes to the PA-induced nonlinear distortion, are on the optimistic side.

Taking into account the input power and maximum gain range of the feeding amplifier, it can also be observed From Table III that the power of the transmitted signal is between -8 and 27 dBm. This is a sufficient range for example in WLAN applications, or in other types of indoor communications. In addition, the studied transmit power range applies in some cases also to mobile devices in a cellular network, like class 3 LTE mobile transmitter [32]. In the following numerical results, the transmit power is varied between -5 and 25 dBm.

B. Results with Case A

In this section, calculations are performed and illustrated under the assumption that the reference signal for RF cancellation is taken after the PA, according to Case A. Thus, the nonlinear distortion produced by the PA is included in the RF reference signal and consequently attenuated by the assumed amount of RF cancellation.

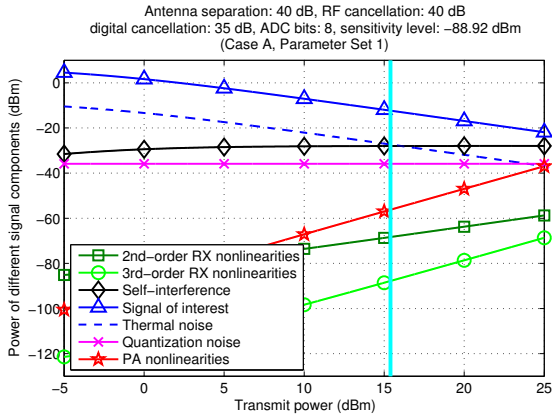


Fig. 3: The power levels of different signal components at the input of the detector with Parameter Set 1, assuming Case A.

1) Fixed Amount of Digital Cancellation: In the first part of the analysis, Parameter Set 1 is used and only the transmit power of the transceiver is varied, while all the other parameters remain constant and unaltered. The power levels of the different signal components can be seen in Fig. 3 in terms of transmit power. The power levels have been calculated using (9)–(16) with the selected parameters. It is imminently obvious that with the chosen parameters, the actual SI is the most significant distortion component. Furthermore, it can be observed that the maximum transmit power is approximately 15 dBm, marked by a vertical line. After this point, the loss of SINR due to SI becomes greater than 3 dB because the SI becomes equally powerful as thermal noise. When interpreting the behavior of the curves in Fig. 3, one should also remember that the power of the signal entering the ADC is kept approximately constant by the AGC. Thus, in practise, the total gain of the RX chain reduces when transmit power increases.

The amount of lost bits, with respect to transmit power, can be seen in Fig. 4. The curve is calculated with (7) and it tells how much of the dynamic range of the ADC is effectively reserved by SI. It can be observed that when using Parameter Set 1, approximately 3 bits are lost due to SI with the maximum transmit power of 15 dBm. This emphasizes the fact that, in this scenario, the actual SI is the limiting factor for the transmit power. Actually, the power of quantization noise is almost 10 dB lower. However, from Fig. 4 it can also be observed that, with a transmit power of 20 dBm, the bit loss is already 4 bits. *This indicates that, in order to enable the usage of higher transmit powers, a high-bit ADC is required.*

2) Variable Amount of Digital Cancellation and Pushing the Performance Limits: In order to further analyze the limits set by the analog-to-digital conversion and nonlinear distortion, it is next assumed that the amount of digital linear cancellation can be increased by an arbitrary amount, while the other parameters remain constant. With this assumption, it is possible to cancel the remaining linear SI perfectly in the digital domain. The reason for performing this type of

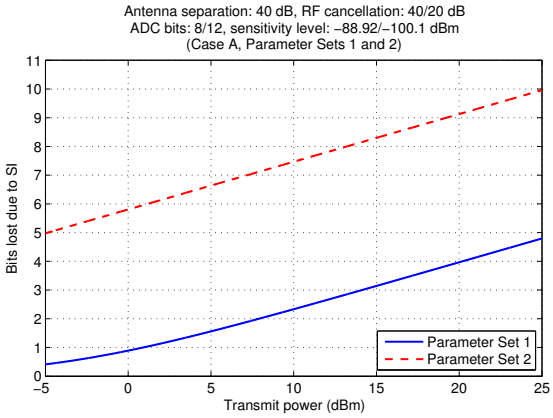


Fig. 4: The amount of lost bits due to SI with both parameter sets in Case A.

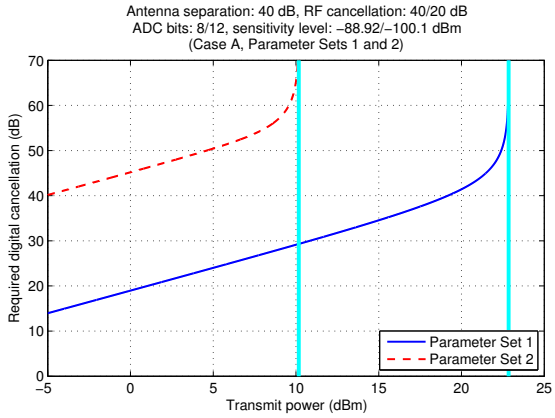


Fig. 5: The required amount of digital cancellation to sustain a 3 dB SINR loss with both parameter sets in Case A.

an analysis is to determine the boundaries of DSP-based SI cancellation, as it would be beneficial to cancel as large amount of SI in the digital domain as possible. However, in many cases, increasing only digital cancellation is not sufficient to guarantee a high enough SINR because nonlinear distortion and quantization noise increase the effective noise floor above the allowed level.

To observe these factors in more detail, the amount of digital cancellation is next selected so that the loss of SINR caused by SI is fixed at 3 dB. This means that the combined power of the other distortion components is allowed to be equal to the power of the thermal noise included in the received signal. Thus, in this case, if the ratio between the signal of interest and dominating distortion becomes smaller than 15 dB, the above condition does not hold, and the loss of SINR becomes greater than 3 dB.

Below we provide closed-form solution for the required amount of digital cancellation. The linear SINR requirement, which must be fulfilled after digital cancellation, is denoted by $sinr_{\mathrm{RQ}}$. Then, the SINR requirement can only be fulfilled if

$$sinr_{\mathrm{RQ}} < \frac{g_{\mathrm{rx}} p_{\mathrm{SOI,in}}}{g_{\mathrm{rx}} F p_{\mathrm{N,in}} + p_{\mathrm{2nd}} + p_{\mathrm{3rd}} + \frac{g_{\mathrm{rx}} p_{\mathrm{3rd,PA,tx}}}{a_{\mathrm{ant}} a_{\mathrm{RF}}} + p_{\mathrm{quant}}}. \quad (19)$$

In words, the SINR must be above the minimum requirement without taking the SI into account. If it is assumed that the above condition holds, the required SINR can be achieved with digital cancellation, and it can be written as

$$sinr_{\mathrm{RQ}} = \frac{g_{\mathrm{rx}} p_{\mathrm{SOI,in}}}{g_{\mathrm{rx}} F p_{\mathrm{N,in}} + \frac{g_{\mathrm{rx}}}{a_{\mathrm{ant}}} \left(\frac{p_{\mathrm{tx}}}{a_{\mathrm{RF}} a_{\mathrm{dig}}} + \frac{p_{\mathrm{3rd,PA,tx}}}{a_{\mathrm{RF}}} \right) + p_{\mathrm{2nd}} + \cdots + p_{\mathrm{3rd}} + p_{\mathrm{quant}}}. \quad (20)$$

From here, the amount of required digital cancellation can be solved and written as

$$\begin{aligned} a_{\mathrm{dig}} &= \frac{\frac{g_{\mathrm{rx}} p_{\mathrm{tx}}}{a_{\mathrm{ant}} a_{\mathrm{RF}}}}{\frac{g_{\mathrm{rx}} p_{\mathrm{SOI,in}}}{sinr_{\mathrm{RQ}}} - (g F p_{\mathrm{N,in}} + p_{\mathrm{2nd}} + p_{\mathrm{3rd}} + \frac{g_{\mathrm{rx}} p_{\mathrm{3rd,PA,tx}}}{a_{\mathrm{ant}} a_{\mathrm{RF}}} + p_{\mathrm{quant}})} \\ &= \frac{1}{1 + \frac{a_{\mathrm{ant}} a_{\mathrm{RF}} p_{\mathrm{SOI,in}}}{p_{\mathrm{tx}}} \left(\frac{1}{sinr_{\mathrm{RQ}}} - \frac{1}{sinr_{\mathrm{DC}}} \right)}, \end{aligned} \quad (21)$$

where $sinr_{\mathrm{DC}}$ is the linear SINR before digital cancellation. The first form of the equation above shows that the amount of required digital cancellation depends directly on the transmit power. It can also be observed that increasing antenna separation or RF cancellation decreases the requirements for digital cancellation.

The required amount of digital cancellation to sustain a maximum of 3 dB SINR loss, calculated from (21), is illustrated in Fig. 5 in terms of the transmit power. The other parameters, apart from digital cancellation, are kept constant. It can be observed from the figure that the maximum transmit power is approximately 23 dBm for Parameter Set 1. After this, the amount of required digital cancellation increases to infinity, indicating perfect linear SI cancellation. However, as discussed earlier, after this point even perfect linear digital cancellation is not sufficient to maintain the required SINR, because quantization noise and nonlinearities become the limiting factor.

The power levels of the different signal components in this scenario are presented in Fig. 6. It can be observed that now quantization noise is the limiting factor for the SINR. The reason for this is that, with higher transmit powers and variable digital cancellation, the majority of SI is now cancelled in the digital domain and thus SI occupies the majority of the dynamic range of the ADC. This, on the other hand, deteriorates the resolution of the desired signal.

In order to further analyze the maximum allowed transmit power of the considered full-duplex transceiver, it is next determined how different parameters influence it. If we denote the signal-to-(thermal)noise-ratio at the detector by snr_{d}, the following equation holds when the loss of SINR is 3 dB:

$$snr_{\mathrm{d}} = \frac{g_{\mathrm{rx}} p_{\mathrm{SOI,in}}}{\frac{g_{\mathrm{rx}}}{a_{\mathrm{ant}}} \left(\frac{p_{\mathrm{tx,max}}}{a_{\mathrm{RF}} a_{\mathrm{dig}}} + \frac{p_{\mathrm{3rd,PA,tx}}}{a_{\mathrm{RF}}} \right) + p_{\mathrm{2nd}} + p_{\mathrm{3rd}} + p_{\mathrm{quant}}}. \quad (22)$$

This means that the total power of the other types of distortion is equal to the thermal noise power, resulting in 3 dB SINR loss. When considering the maximum transmit power, it is again assumed that digital SI cancellation is perfect. Furthermore, as the transmit power is high, and also the nonlinear

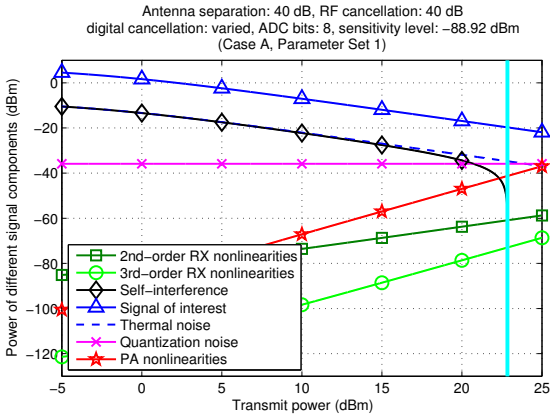


Fig. 6: The power levels of different signal components at the input of the detector when the amount of digital cancellation is increased, assuming Parameter Set 1 and Case A.

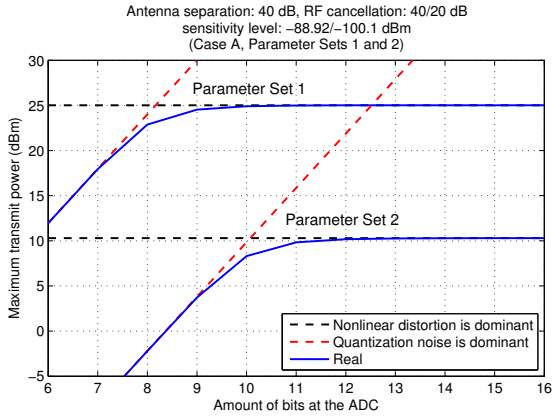


Fig. 7: The maximum transmit power with respect to the number of bits at the ADC, again with both parameter sets. The blue curve shows the real value of the maximum transmit power, and the red and black curves show the values when quantization noise or nonlinear distortion is the dominant distortion, respectively.

distortion produced by the PA is attenuated by RF cancellation, the power of SI can be used to approximate the power of the total signal at the input of the RX chain. This, on the other hand, allows us to use (11) to approximate the total receiver gain. Thus, when substituting g_{rx} with (11), letting $a_{\text{dig}} \to \infty$, and expressing quantization noise as $\frac{p_{\text{target}}}{snr_{\text{ADC}}}$, (22) becomes

$$\begin{aligned} snr_{\text{d}} &= \frac{\frac{a_{\text{ant}}a_{\text{RF}}p_{\text{target}}}{p_{\text{tx,max}}}p_{\text{SOI,in}}}{\frac{a_{\text{ant}}a_{\text{RF}}p_{\text{target}}}{p_{\text{tx,max}}}\frac{p_{\text{3rd,PA,tx}}}{a_{\text{ant}}a_{\text{RF}}} + p_{\text{2nd}} + p_{\text{3rd}} + \frac{p_{\text{target}}}{snr_{\text{ADC}}}} \\ &= \frac{a_{\text{ant}}a_{\text{RF}}p_{\text{SOI,in}}}{p_{\text{tx,max}}\left(\frac{p_{\text{2nd}}+p_{\text{3rd}}}{p_{\text{target}}} + \frac{1}{snr_{\text{ADC}}}\right) + p_{\text{3rd,PA,tx}}}. \end{aligned} \tag{23}$$

By solving (23) in terms of $p_{\text{tx,max}}$, the maximum transmit power can be obtained. However, as the power of nonlinear distortion is dependent on the transmit power, it is not convenient to derive an analytical equation for the maximum transmit power as it would require solving the roots of a 3rd-order polynomial. On the other hand, if we consider the scenario of Fig. 6, it can be seen that the quantization noise is actually the dominant distortion component. Thus, in this case, $p_{\text{2nd}} + p_{\text{3rd}} \approx 0$ and $p_{\text{3rd,PA,tx}} \approx 0$, and the maximum transmit power becomes

$$\begin{aligned} p_{\text{tx,max}} &= \frac{a_{\text{ant}}a_{\text{RF}}p_{\text{SOI,in}}snr_{\text{ADC}}}{snr_{\text{d}}} \\ \text{i.e. } P_{\text{tx,max}} &= A_{\text{ant}} + A_{\text{RF}} + P_{\text{SOI,in}} + SNR_{\text{ADC}} - SNR_{\text{d}}. \end{aligned} \tag{24}$$

By substituting SNR_{ADC} with (5), we can approximate the maximum transmit power of the considered full-duplex transceiver as

$$\begin{aligned} P_{\text{tx,max}} = A_{\text{ant}} + A_{\text{RF}} + P_{\text{SOI,in}} - SNR_{\text{d}} + 6.02b \\ - PAPR + 4.76. \end{aligned} \tag{25}$$

This applies accurately when the quantization noise is the limiting factor.

An alternative possible scenario is the situation where the amount of bits is sufficiently high such that the quantization noise is not the main performance bottleneck. In this case, the power of nonlinear distortion is the limiting factor for the maximum transmit power (still assuming $a_{\text{dig}} \to \infty$). In other words, if we let $snr_{\text{ADC}} \to \infty$, (23) becomes

$$snr_{\text{d}} = \frac{a_{\text{ant}}a_{\text{RF}}p_{\text{SOI,in}}}{p_{\text{tx,max}}\left(\frac{p_{\text{2nd}}+p_{\text{3rd}}}{p_{\text{target}}}\right) + p_{\text{3rd,PA,tx}}}. \tag{26}$$

However, similar to solving (23), it is again very inconvenient to derive a compact form for the maximum transmit power in this scenario, since it would again require solving the roots of a third order polynomial. Nevertheless, the value for the maximum transmit power can in this case be easily calculated numerically, which yields $p_{\text{tx,max}} \approx 25.02$ dBm and $p_{\text{tx,max}} \approx 10.29$ dBm with Parameter Sets 1 and 2, respectively.

If operating under such conditions that neither intermodulation nor quantization noise is clearly dominating, previous results in (25) and (26) may be overestimating the performance. For this reason, Fig. 7 shows the actual maximum transmit power with respect to the number of bits at the ADC without any such assumptions, calculated numerically from (23). Also the maximum transmit powers for the two special scenarios are shown. With a low number of bits, the quantization noise is indeed the limiting factor for the transmit power, and the curve corresponding to (25) is very close to the real value. On the other hand, with a high number of bits, the line corresponding to (26) is closer to the real value, as the power of quantization noise becomes negligibly low. This demonstrates very good accuracy and applicability of the derived analytical results.

Perhaps the most interesting observation from Fig. 7 is that, with Parameter Set 1, it is sufficient to have a 10 bit ADC in order to decrease the power of quantization noise negligibly low. This is shown by the fact that after that point, the maximum transmit power saturates to the value calculated with (26). The saturated value of the maximum transmit power can only be increased by implementing more linear transceiver components or by increasing the amount of SI attenuation in the analog domain, thereby decreasing the power of nonlinear distortion and thus lowering the overall noise floor.

Overall, with the chosen parameters for the receiver, the bottleneck during the full-duplex operation in Case A is the quantization noise, in addition to the actual SI. This is an observation worth noting, as performing as much SI cancellation in the digital domain as possible is very desirable, since it allows the construction of cheaper and more compact full-duplex transceivers with affordable and highly-integrated RF components. In addition, it is also observed that, with higher transmit powers, the nonlinear distortion produced by the PA of the transmitter is a considerable factor. If a cheaper and less linear PA is used, this nonlinear distortion starts to limit even more heavily the achievable performance of a full-duplex transceiver.

3) Calculations with Parameter Set 2: In order to analyze how using cheaper, and hence lower-quality, components affects the RX chain, some calculations are done also with Parameter Set 2. The values of the parameters are again listed in Tables I and II. The sensitivity of the receiver is improved by decreasing the bandwidth and SNR requirement, and the power of the received signal is also decreased accordingly. In addition, the amount of RF cancellation is now assumed to be only 20 dB. This has a serious effect on the bit loss and the requirements for the digital cancellation.

The only component, whose specifications are improved, is the ADC, as it is now chosen to have 12 bits. The reason for this is to preserve a sufficient resolution for the signal of interest in the digital domain, as the amount of lost bits is relatively high with these weaker parameters. The calculations are again carried out assuming that the amount of digital cancellation can be increased arbitrarily high.

The required amount of digital cancellation, when using Parameter Set 2, is depicted in Fig. 5, and Fig. 8 shows the power levels of the different signal components in this scenario, again calculated with (9)–(18). It can be seen that now nonlinear distortion, produced by the receiver components, is the limiting factor for the transmit power, instead of quantization noise. The maximum transmit power is only approximately 10 dBm. After this point, mitigating only the linear SI is not sufficient to sustain the required SINR, as nonlinear distortion decreases the SINR below the required level.

With this parameter set, it can be seen that the amount of lost bits is very high (cf. Fig. 4). This is due to the decreased RF cancellation ability, which means that the SI power is higher at the ADC interface. Thus, with lower SI cancellation performance at the analog/RF domain, the requirements for the ADC are heavily increased.

It can also be concluded that, with cheaper and less linear components, mitigating the RX chain nonlinear distortion with additional nonlinear DSP can provide performance gain. This is shown by Fig. 7, where it can be observed that with Parameter Set 2, the maximum transmit power is decreased to 10 dBm, as opposed to the maximum transmit power of 25 dBm achieved with Parameter Set 1. This difference is caused by the lower linearity and decreased RF cancellation ability of the receiver utilizing Parameter Set 2. Thus, with decreased transceiver linearity and RF cancellation ability, also the nonlinear distortion produced by the RX chain must be considered, as Figs. 7 and 8 demonstrate.

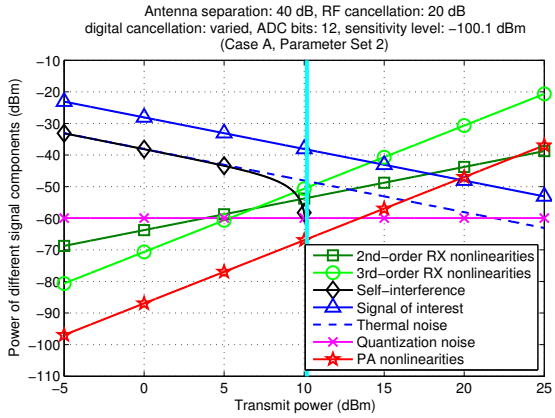


Fig. 8: The power levels of different signal components at the input of the detector with Parameter Set 2, assuming Case A.

C. Results with Case B

In the system calculations of this section, Case B is considered, and thus the reference signal for RF cancellation is taken from the input of the PA. This means that the nonlinear distortion produced by the PA is not attenuated by RF cancellation, as it is not included in the cancellation signal. This obviously increases the effect of these TX-induced nonlinearities.

The values for the parameters of the RX chain are chosen according to Parameter Set 1, and the amount of digital cancellation is again controlled to maintain a 3 dB loss of SINR. The transmit power is varied from -5 dBm to 25 dBm. Figure 9 illustrates the power levels of different signal components in this scenario. It can be observed that the nonlinear distortion produced by the PA is the most significant distortion component already with transmit powers higher than 11 dBm. Furthermore, with transmit powers higher than 12 dBm, it will decrease the SINR below the required level, thus preventing the usage of higher transmit powers.

When comparing Fig. 9 to Fig. 6, it can be observed that the difference is significant. This is caused by the fact that in Case B, the nonlinear distortion produced by the PA is not attenuated by RF cancellation, unlike in Case A. Hence, it is clear that an ability to mitigate nonlinear distortion would provide a significant performance gain for a full-duplex transceiver, which is implemented according to Case B. Furthermore, with the chosen parameters, it would be sufficient to mitigate the nonlinearities in the digital domain, as the quantization noise floor is fairly low relative to the other signal components.

In order to demonstrate the potential of nonlinear cancellation, the maximum transmit powers of two different scenarios are compared. In the first case, it is assumed that digital cancellation is linear, and can thus mitigate only the linear part of the SI signal. In the other case, it is assumed that digital cancellation is able to mitigate also the nonlinear part of the SI signal, in addition to the linear part. Figure 10 shows the increase in the maximum transmit power, when comparing these two scenarios. The same curve has also been plotted

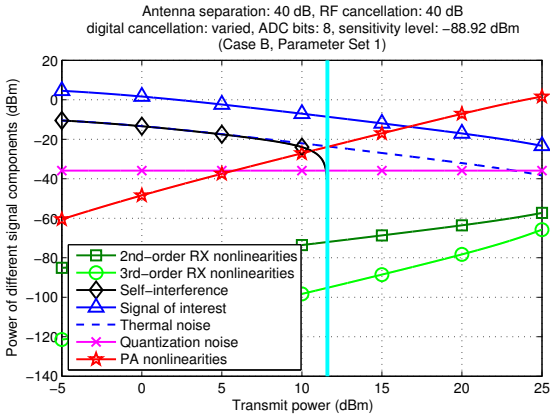


Fig. 9: The power levels of different signal components at the input of the detector with Parameter Set 1, assuming Case B.

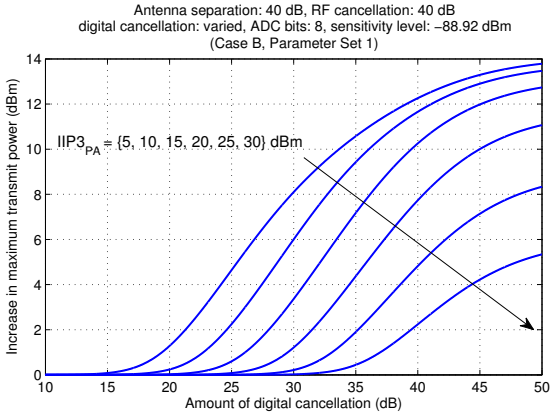


Fig. 10: The increase in maximum transmit power when also the nonlinear distortion of the SI channel can be mitigated with digital cancellation, compared to only linear cancellation. Horizontal axis depicts the total amount of achieved digital cancellation. The curves correspond to different IIP3 figures of the PA.

with different IIP3 values for the PA. The curves have been calculated based on (9)–(18), with the modification that in the other case, $P_{\mathrm{3rd,PA}}$ is also attenuated by A_{dig}. It can be observed that being able to mitigate the nonlinear component of the SI signal in the digital domain provides a significant increase in the maximum transmit power when the total amount of digital cancellation is increased. This has also been observed with actual waveform simulations in [15].

It can also be observed that already with 25 dB of digital cancellation, the maximum transmit power is increased by as much as 5 dB, if also the nonlinear component of the SI signal is mitigated. Obviously, the achievable gain is smaller with a more linear PA, and this indicates that when the nonlinear component of the SI signal is weaker, linear digital cancellation might be sufficient. However, with a less linear PA, significant increase in the maximum transmit power can be achieved with nonlinear digital cancellation, almost regardless of the total amount of achieved cancellation.

TABLE IV: Additional parameters for the waveform simulator.

Parameter	Value
Constellation	16-QAM
Number of subcarriers	64
Number of data subcarriers	48
Guard interval	16 samples
Sample length	15.625 ns
Symbol length	4 μs
Signal bandwidth	12.5 MHz
Oversampling factor	4
ADC bits	12

Overall, Figs. 9 and 10 illustrate that nonlinear distortion produced by the TX PA is a significant issue in full-duplex transceivers, when the reference signal for RF cancellation is taken from the input of the PA. Furthermore, the ability to compensate it can significantly improve the performance of the transceiver. Thus, implementing nonlinear estimation and processing mechanisms for digital SI cancellation is an interesting topic for future research.

IV. Waveform Simulations and Comparisons

In order to analyze and demonstrate the good accuracy of the used models and the system calculation results, a complete full-duplex waveform simulator is constructed. It emulates a similar direct conversion transceiver that is used in the analytical calculations, having the parameters corresponding to Parameter Set 1. Here, only Case A is considered for compactness.

The simulator is implemented with Matlab and Simulink, using SimRF component library. The simulated waveform is chosen to be an OFDM signal with parameters specified in Table IV. These parameters are in essence similar to WLAN specifications, and they are used for generating both the transmitted and received signals.

The SI channel is assumed to be static and it consists of a main coupling component and three weak multipath components, which are delayed by one, three, and eight sample intervals in relation to the main component, respectively. This corresponds to a maximum delay of 125 ns. The delay of the main component is assumed to be negligibly small, as the distance between the antennas is typically very short. The average power difference between the main component and the multipath components is set to 45 dB, which is on the same range as values measured in [29].

In these simulations, a single-tap RF canceller is considered, as this corresponds to the most typical scenario currently used in the literature [3], [4]. Thus, in this scenario, RF cancellation attenuates only the main coupling component of the SI signal. In the simulator, also some delay, amplitude, and phase errors are included in the RF cancellation signal to achieve the desired amount of SI attenuation, and to model the cancellation process in a realistic manner.

The attenuation of the weaker multipath components is then done by digital cancellation after the ADC. The implementation of digital cancellation utilizes classic least-squares based

SI coupling channel estimation, which is implemented with linear least-squares fitting between the ideal TX data and RX observation during a calibration period. Thus, the amount of digital cancellation cannot be tuned arbitrarily since it depends directly on the accuracy of these TX-RX channel estimates. The amount of achieved digital cancellation is illustrated in Fig. 11. The fluctuating curve is the realized value, and the smooth curve is a third order polynomial fitted to the realized values. The polynomial approximation is used when calculating the analytical SINR, in order to assess realistic average performance. As shown in Fig. 11, larger amounts of cancellation are achieved with higher transmit powers, as the quality of the channel estimate is better with a stronger SI signal. This phenomenon has also been observed in practice [29]. However, with transmit powers above 17 dBm, the power of the PA-induced nonlinear distortion starts to decrease the achievable digital cancellation.

The results of the analytical calculations are then compared to the simulation results in terms of the SINR at the input of the detector ($SINR_{\mathrm{d}}$). Figure 12 shows the SINRs obtained with analytical calculations and with full waveform simulations, with respect to transmit power. In the waveform simulator, the SINR is calculated by first determining the effective powers for the ideal signal, and total noise-plus-interference signal. After this, the SINR is calculated as the ratio of these signal powers. The simulation is repeated 50 times for each transmit power, and the transmit power is varied with 1 dB intervals. The SINR corresponding to each transmit power is calculated as the average value of these independent realizations. The analytical SINR is calculated directly from the previously presented equations. From Fig. 12 it can be seen that the analytical and simulated SINR curves are practically identical, thus evidencing excellent accuracy and reliability of the reported analytical expressions. With closer inspection, it can be observed that the analytically calculated SINR is actually slightly pessimistic throughout the considered transmit power range, but the difference is only in the order of 0.1–0.3 dB. This is likely to be caused by the different approximations made when deriving the equations for the power levels of the different signal components. In any case, it can be concluded that the accuracy of the analysis is very high.

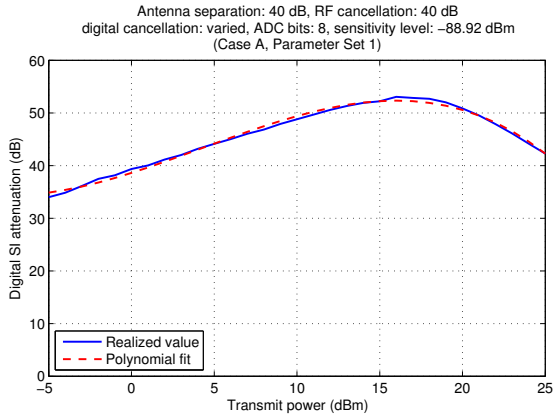


Fig. 11: The amount of achieved digital cancellation in the true waveform simulation, with respect to transmit power. TX-RX channel estimation in the digital cancellation is implemented with linear least-squares fitting between the ideal TX data and RX observation during a calibration period.

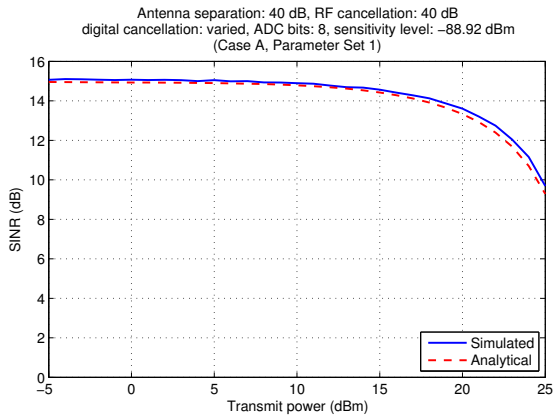


Fig. 12: SINR values obtained from the waveform simulations and from the analytical calculations.

V. Conclusion

In this paper, the effects of self-interference on the receiver chain of a full-duplex transceiver were analyzed in detail, taking into accout realistic antenna isolation, RF cancellation and digital baseband cancellation. Specific emphasis was on modeling and analyzing the impacts of transmitter and receiver RF nonlinearities as well as analog-to-digital converter dynamic range requirements. The reliability of the analytical results was also verified and demonstrated by comparing them with the reference results acquired from complete full-duplex device waveform simulations. These comparisons showed excellent match, verifying the high accuracy and reliability of the results.

In terms of RF cancellation reference injection, the analysis covered two alternative scenarios where the reference is taken either from transmitter power amplifier output (Case A) or input (Case B). In Case A, it was observed that with high-quality RF components and RF cancellation, the maximum tolerable transmit power is mostly limited by the quantization noise of the receiver analog-to-digital converter, as well as by the achievable amount of linear digital cancellation. However, with low-cost receiver RF components and lower-quality RF cancellation, feasible for mobile devices, also the transceiver chain nonlinearities were found considerable and can actually become the limiting factor. In Case B, in turn, it was observed that the linearity of the power amplifier is the major bottleneck for the receiver performance with transmit powers above 10 dBm, even when clearly fulfilling any typical transmitter emission mask. This applies also to a closely-related architecture where separate low-power transmitter chain is used to generate the RF reference.

In order to be able to implement full-duplex transceivers with transmit powers in the order of 20–30 dBm, typical to

WiFi and mobile cellular radio terminals, with low-cost RF electronics, the findings of this article strongly motivate for the development of nonlinear digital self-interference cancellation techniques. This applies to the 2nd- and 3rd-order inband nonlinear distortion of the receiver RF components, and in particular to the 3rd-order inband nonlinear distortion of the transmitter power amplifier. Developing such nonlinear cancellation techniques is the main topic of our future research.

APPENDIX A
DERIVATION OF BIT LOSS DUE TO SI

A principal equation for the bit loss due to noise and interference is written in (6). However, as we are now interested in the amount of bits lost due to SI, the bit losses under HD and FD operation must be compared. By subtracting the amount of lost bits under HD operation from the amount of lost bits under FD operation, we obtain the desired value of bit loss due to SI, written as

$$b_{\text{lost}} = \frac{P_{\text{target}} - P_{\text{SOI,FD}}}{6.02} - \frac{P_{\text{target}} - P_{\text{SOI,HD}}}{6.02}, \tag{27}$$

where P_{target} corresponds to the total power of the signal at the input of the ADC (which is always constant because of AGC), and $P_{\text{SOI,FD}}$ and $P_{\text{SOI,HD}}$ are the powers of the desired signal with and without SI, respectively. Because the total power of the signal at the input of the ADC is kept constant by the AGC, (27) can be further simplified to express the bit loss in terms of the gains as

$$\begin{aligned} b_{\text{lost}} &= \frac{P_{\text{SOI,HD}} - P_{\text{SOI,FD}}}{6.02} \\ &= \frac{P_{\text{SOI,in}} + G_{\text{HD}} - (P_{\text{SOI,in}} + G_{\text{FD}})}{6.02} = \frac{G_{\text{HD}} - G_{\text{FD}}}{6.02}, \end{aligned} \tag{28}$$

where G_{FD} is the total gain of the RX chain under FD operation, and G_{HD} is the total gain under HD operation, correspondingly. This is a rather intuitive expression for the bit loss, as the power of SI is obviously included in G_{FD} due the reduction of the gain by the AGC. Noting that $G = P_{\text{target}} - P_{\text{in}}$ and $6.02 \approx 10\log_{10}(4)$, the bit loss can be written as

$$\begin{aligned} b_{\text{lost}} &= \frac{(P_{\text{target}} - P_{\text{in,HD}}) - (P_{\text{target}} - P_{\text{in,FD}})}{10\log_{10}(4)} \\ &= \frac{P_{\text{in,FD}} - P_{\text{in,HD}}}{10\log_{10}(4)} = \frac{10\log_{10}\left(\frac{p_{\text{in,FD}}}{p_{\text{in,HD}}}\right)}{10\log_{10}(4)} = \log_4\left(\frac{p_{\text{in,FD}}}{p_{\text{in,HD}}}\right) \\ &\approx \log_4\left(1 + \frac{p_{\text{SI,in}} + p_{\text{3rd,PA,in}}}{p_{\text{SOI,in}} + p_{\text{N,in}}}\right). \end{aligned} \tag{29}$$

By denoting that $p_{\text{SI,in}} = \frac{p_{\text{tx}}}{a_{\text{ant}}a_{\text{RF}}}$ and $p_{\text{3rd,PA,in}} = \frac{p_{\text{3rd,PA,tx}}}{a_{\text{ant}}a_{\text{NL}}} = \frac{p_{\text{tx}}^3}{a_{\text{ant}}a_{\text{NL}}iip3_{\text{PA}}^2 g_{\text{PA}}^2}$, (29) can finally be written as

$$\begin{aligned} b_{\text{lost}} = \log_4\Bigg[1 + \left(\frac{1}{p_{\text{SOI,in}} + p_{\text{N,in}}}\right) \\ \cdot \left(\frac{p_{\text{tx}}}{a_{\text{ant}}a_{\text{RF}}} + \frac{p_{\text{tx}}^3}{a_{\text{ant}}a_{\text{NL}}iip3_{\text{PA}}^2 g_{\text{PA}}^2}\right)\Bigg]. \end{aligned} \tag{30}$$

APPENDIX B
DERIVATIONS OF RECEIVER NONLINEAR DISTORTION PRODUCTS

The derivation of (15) and (16) is done based on the power of nonlinear distortion at the output of a single component. This, on the other hand, can be calculated with (8). In the considered full-duplex transceiver, only the mixer and the VGA produce SI-induced 2nd-order nonlinear distortion on to the signal band. However, all the components are assumed to produce 3rd-order nonlinear distortion.

The derivation is done with linear power units to present the calculations in a more compact form. The total power of the signal at the input of the RX chain is denoted as p_{in}. It consists of the signal of interest, SI, and thermal noise. Furthermore, the increase in the thermal noise power occurring within the RX chain is omitted, as it has no significant effect on the power of the nonlinear distortion. Using (8), and expressing the output power in terms of gain and input power, the power of the 3rd-order nonlinear distortion at the output of the LNA can be written as

$$P_{\text{3rd,LNA}} = G_{\text{LNA}} + P_{\text{in}} - 2(IIP3_{\text{LNA}} - P_{\text{in}}) \tag{31}$$

Using the corresponding linear units, this can be written as

$$p_{\text{3rd,LNA}} = \frac{g_{\text{LNA}}p_{\text{in}}^3}{iip3_{\text{LNA}}^2}. \tag{32}$$

Now, noting that with the chosen parameters the power of the nonlinear distortion is negligibly small in comparison to the total power of the signal, the input power of the mixer can be written as

$$p_{\text{in,mixer}} = g_{\text{LNA}}p_{\text{in}} + p_{\text{3rd,LNA}} \approx g_{\text{LNA}}p_{\text{in}}. \tag{33}$$

The power of the 2nd-order nonlinear distortion produced by the mixer can be then written as

$$\begin{aligned} p_{\text{2nd,mixer}} &= \frac{g_{\text{mixer}}p_{\text{in,mixer}}^2}{iip2_{\text{mixer}}} = \frac{g_{\text{mixer}}}{iip2_{\text{mixer}}}\left(g_{\text{LNA}}p_{\text{in}}\right)^2 \\ &= \frac{g_{\text{LNA}}^2 g_{\text{mixer}}p_{\text{in}}^2}{iip2_{\text{mixer}}}. \end{aligned} \tag{34}$$

The power of the 3rd-order nonlinear distortion produced by the mixer can in turn be written as

$$\begin{aligned} p_{\text{3rd,mixer}} &= \frac{g_{\text{mixer}}p_{\text{in,mixer}}^3}{iip3_{\text{mixer}}^2} = \frac{g_{\text{mixer}}}{iip3_{\text{mixer}}^2}\left(g_{\text{LNA}}p_{\text{in}}\right)^3 \\ &= \frac{g_{\text{LNA}}^3 g_{\text{mixer}}p_{\text{in}}^3}{iip3_{\text{mixer}}^2}. \end{aligned} \tag{35}$$

Again, noting that the power of the nonlinear distortion is negligibly small in comparison to the total power of the signal, the input power of the VGA can be written as

$$p_{\text{in,VGA}} \approx g_{\text{mixer}}p_{\text{in,mixer}} = g_{\text{LNA}}g_{\text{mixer}}p_{\text{in}}. \tag{36}$$

The power of the 2nd-order nonlinear distortion at the output of the VGA can thus be written as

$$\begin{aligned} p_{\text{2nd,VGA}} &= \frac{g_{\text{VGA}}p_{\text{in,VGA}}^2}{iip2_{\text{VGA}}} = \frac{g_{\text{VGA}}}{iip2_{\text{VGA}}}\left(g_{\text{LNA}}g_{\text{mixer}}p_{\text{in}}\right)^2 \\ &= \frac{g_{\text{LNA}}^2 g_{\text{mixer}}^2 g_{\text{VGA}}p_{\text{in}}^2}{iip2_{\text{VGA}}}. \end{aligned} \tag{37}$$

Similarly, the power of 3rd-order nonlinear distortion at the output of the VGA can be written as

$$p_{\text{3rd,VGA}} = \frac{g_{\text{VGA}} p_{\text{in,VGA}}^3}{iip3_{\text{VGA}}^2} = \frac{g_{\text{VGA}}}{iip3_{\text{VGA}}^2} \left(g_{\text{LNA}} g_{\text{mixer}} p_{\text{in}}\right)^3 = \frac{g_{\text{LNA}}^3 g_{\text{mixer}}^3 g_{\text{VGA}} p_{\text{in}}^3}{iip3_{\text{VGA}}^2}. \quad (38)$$

Finally, the total power of the nonlinear distortion of a given order can be determined by summing the powers of the nonlinear distortion at the output of each individual component, including also the effect of the gains of the upcoming components. Thus, the total power of the 2nd-order nonlinear distortion can be written as follows, using (34) and (37):

$$\begin{aligned} p_{\text{2nd}} &= g_{\text{VGA}} p_{\text{2nd,mixer}} + p_{\text{2nd,VGA}} \\ &= g_{\text{VGA}} \frac{g_{\text{LNA}}^2 g_{\text{mixer}} p_{\text{in}}^2}{iip2_{\text{mixer}}} + \frac{g_{\text{LNA}}^2 g_{\text{mixer}}^2 g_{\text{VGA}} p_{\text{in}}^2}{iip2_{\text{VGA}}} \\ &= g_{\text{LNA}}^2 g_{\text{mixer}} g_{\text{VGA}} p_{\text{in}}^2 \left(\frac{1}{iip2_{\text{mixer}}} + \frac{g_{\text{mixer}}}{iip2_{\text{VGA}}} \right). \end{aligned} \quad (39)$$

Similarly, the total power of the 3rd-order nonlinear distortion can be written as follows, using (32), (35), and (38):

$$\begin{aligned} p_{\text{3rd}} &= g_{\text{mixer}} g_{\text{VGA}} p_{\text{3rd,LNA}} + g_{\text{VGA}} p_{\text{3rd,mixer}} + p_{\text{3rd,VGA}} \\ &= g_{\text{mixer}} g_{\text{VGA}} \frac{g_{\text{LNA}} p_{\text{in}}^3}{iip3_{\text{LNA}}^2} + g_{\text{VGA}} \frac{g_{\text{LNA}}^3 g_{\text{mixer}} p_{\text{in}}^3}{iip3_{\text{mixer}}^2} \\ &\quad + \frac{g_{\text{LNA}}^3 g_{\text{mixer}}^3 g_{\text{VGA}} p_{\text{in}}^3}{iip3_{\text{VGA}}^2} \\ &= g_{\text{LNA}} g_{\text{mixer}} g_{\text{VGA}} p_{\text{in}}^3 \left[\left(\frac{1}{iip3_{\text{LNA}}} \right)^2 \right. \\ &\quad \left. + \left(\frac{g_{\text{LNA}}}{iip3_{\text{mixer}}} \right)^2 + \left(\frac{g_{\text{LNA}} g_{\text{mixer}}}{iip3_{\text{VGA}}} \right)^2 \right]. \end{aligned} \quad (40)$$

When comparing the values calculated with the obtained equations to the values calculated without approximations, it is observed that the error is in the order of 0.7 % with transmit powers above 5 dBm. Thus, the approximations which are made in the derivation process do not have any notable effect on the reliability of the equations with the chosen parameter range.

REFERENCES

[1] M. E. Knox, "Single antenna full duplex communications using a common carrier," in *Proc. 13th Annual IEEE Wireless and Microwave Technology Conference*, Apr. 2012, pp. 1–6.

[2] C. Cox and E. Ackerman, "Demonstration of a single-aperture, full-duplex communication system," in *Proc. Radio and Wireless Symposium*, Jan. 2013, pp. 148–150.

[3] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel full duplex wireless communication," in *Proc. 16th Annual International Conference on Mobile Computing and Networking*, Sep. 2010, pp. 1–12.

[4] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. 17th Annual International Conference on Mobile computing and Networking*, Jan. 2011, pp. 301–312.

[5] A. Sahai, G. Patel, and A. Sabharwal, "Pushing the limits of full-duplex: Design and real-time implementation," Department of Electrical and Computer Engineering, Rice University, Technical Report TREE1104, Jul. 2011.

[6] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. 44th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2010, pp. 1558–1562.

[7] T. Riihonen, P. Mathecken, and R. Wichman, "Effect of oscillator phase noise and processing delay in full-duplex OFDM repeaters," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 1947–1951.

[8] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "Understanding the impact of phase noise on active cancellation in wireless full-duplex," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 29–33.

[9] V. Syrjälä, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," *to appear in IEEE Transactions on Wireless Communications*, 2014. [Online]. Available: www.cs.tut.fi/~syrjalav/publications/FD_pn.pdf

[10] E. Ahmed, A. Eltawil, and A. Sabharwal, "Self-interference cancellation with phase noise induced ICI suppression for full-duplex systems," *CoRR*, vol. abs/1307.4149, 2013. [Online]. Available: http://arxiv.org/abs/1307.4149

[11] A. Sahai, G. Patel, and A. Sabharwal, "Asynchronous full-duplex wireless," in *Proc. Fourth International Conference on Communication Systems and Networks (COMSNETS)*, Jan. 2012, pp. 1–9.

[12] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely-linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *accepted with minor revisions to IEEE JSAC special issue on Full-Duplex Wireless Communications and Networks*, to be published, 2014. [Online]. Available: http://arxiv.org/abs/1402.6083

[13] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *SIGCOMM'13*, Aug. 2013.

[14] E. Ahmed, A. Eltawil, and A. Sabharwal, "Self-interference cancellation with nonlinear distortion suppression for full-duplex systems," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*, Nov. 2013.

[15] L. Anttila, D. Korpi, V. Syrjälä, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*, Nov. 2013. [Online]. Available: http://arxiv.org/abs/1401.3129

[16] D. Korpi et al., "Advanced self-interference cancellation and multiantenna techniques for full-duplex radios," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*, Nov. 2013. [Online]. Available: http://arxiv.org/abs/1401.3331

[17] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2498–2511, May 2013.

[18] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[19] B. Day, A. Margetts, D. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3702–3713, Jul. 2012.

[20] ——, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 8, pp. 1541–1553, Sep. 2012.

[21] T. Riihonen and R. Wichman, "Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 45–49.

[22] Q. Gu, *RF System Design of Transceivers for Wireless Communications*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[23] B. Razavi, *RF microelectronics*, ser. Prentice Hall communications engineering and emerging technologies series. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1998.

[24] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.

[25] S. Li and R. Murch, "Full-duplex wireless communication using transmitter output based echo cancellation," in *Proc. Global Telecommunications Conference*, Dec. 2011, pp. 1–5.

[26] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, V. Balan, and D. Gerald, "Rethinking indoor wireless: Low power, low frequency, full-duplex," Microsoft Research, Microsoft Corporation, Technical Report MSR-TR-2009-148, 2009.

[27] S. Chen and M. Beach, "Division-free duplex for wireless applications," *Electronics Letters*, vol. 34, no. 2, pp. 147–148, Jan. 1998.

[28] "QHx-220 active isolation enhancer and interference canceller," Intersil, Milpitas, California, USA.
[29] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.
[30] F. Raab, P. Asbeck, S. Cripps, P. Kenington, Z. Popovic, N. Pothecary, J. Sevic, and N. Sokal, "Power amplifiers and transmitters for RF and microwave," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 814–826, Mar. 2002.
[31] E. Ahmed, A. Eltawil, and A. Sabharwal, "Rate gain region and design tradeoffs for full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3556–3565, Jul. 2013.
[32] "LTE; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception (3GPP TS 36.101 version 11.2.0 release 11)," ETSI, Sophia Antipolis Cedex, France.
[33] H. Yoshida, T. Kato, T. Toyoda, I. Seto, R. Fujimoto, T. Kimura, O. Watanabe, T. Arai, T. Itakura, and H. Tsurumi, "Fully differential direct conversion receiver for W-CDMA using an active harmonic mixer," in *Proc. IEEE Radio Frequency Integrated Circuits Symposium*, Jun. 2003, pp. 395–398.
[34] A. Pärssinen, J. Jussila, J. Ryynänen, L. Sumanen, and K. A. I. Halonen, "A 2-GHz wide-band direct conversion receiver for WCDMA applications," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 12, pp. 1893–1903, Dec. 1999.
[35] A. Behzad, *Wireless LAN Radios: System Definition to Transistor Design*, ser. IEEE Series on Digital and Mobile Communication. Wiley, 2007.
[36] "TLV1548-Q1 low-voltage 10-bit analog-to-digital converter with serial control and 8 analog inputs," Texas Instruments, Dallas, Texas, USA.

Dani Korpi was born in Ilmajoki, Finland, on November 16, 1989. He received the B.Sc. degree (with honors) in communications engineering from Tampere University of Technology (TUT), Tampere, Finland, in 2012, and is currently pursuing the M.Sc. degree in communications engineering at TUT.

In 2011, he was a Research Assistant with the Department of Signal Processing at TUT. Since 2012, he has been a Research Assistant with the Department of Electronics and Communications Engineering, TUT. His main research interest is the study and development of single-channel full-duplex radios, with a focus on analysing the RF impairments.

Taneli Riihonen (S'06, M'14) received the M.Sc. degree in communications engineering (with distinction) from Helsinki University of Technology, Helsinki, Finland in February 2006.

During the summer of 2005, he was an intern at Nokia Research Center, Helsinki, Finland. Since fall 2005, he has been a researcher at the Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, Helsinki, Finland, where he is completing his D.Sc. (Tech.) degree in the near future. He has also been a student at the Graduate School in Electronics, Telecommunications and Automation (GETA) in 2006-2010. His research activity is focused on physical-layer OFDM(A), multiantenna and relaying techniques.

Ville Syrjälä (S'09, M'12) was born in Lapua, Finland, in 1982. He received the M.Sc. (Tech.) degree in 2007 and D.Sc. (Tech.) degree in 2012 in communications engineering (CS/EE) from Tampere University of Technology (TUT), Finland.

He was working as a research fellow with the Department of Electronics and Communications Engineering at TUT, Finland, until 2013. Currently, he is working as a research fellow of the Japan Society for the Promotion of Science (JSPS) at Kyoto University, Japan. His general research interests are in full-duplex radio technology, communications signal processing, transceiver impairments, signal processing algorithms for flexible radios, transceiver architectures, direct sampling radios, and multicarrier modulation techniques.

Lauri Anttila (S'06, M'11) received his Ph.D. (with honours) from Tampere University of Technology (TUT), Tampere, Finland, in 2011.

Currently, he is a Research Fellow at the Department of Electronics and Communications Engineering at TUT. His current research interests include statistical and adaptive signal processing for communications, digital front-end signal processing in flexible radio transceivers, and full-duplex radio systems.

Mikko Valkama (S'00, M'02) was born in Pirkkala, Finland, on November 27, 1975. He received the M.Sc. and Ph.D. degrees (both with honours) in electrical engineering (EE) from Tampere University of Technology (TUT), Finland, in 2000 and 2001, respectively. In 2002 he received the Best Ph.D. Thesis award by the Finnish Academy of Science and Letters for his dissertation entitled "Advanced I/Q signal processing for wideband receivers: Models and algorithms".

In 2003, he was working as a visiting researcher with the Communications Systems and Signal Processing Institute at SDSU, San Diego, CA. Currently, he is a Full Professor and Department Vice Head at the Department of Electronics and Communications Engineering at TUT, Finland. He has been involved in organizing conferences, like the IEEE SPAWC07 (Publications Chair) held in Helsinki, Finland. His general research interests include communications signal processing, estimation and detection techniques, signal processing algorithms for software defined flexible radios, full-duplex radio technology, cognitive radio, digital transmission techniques such as different variants of multicarrier modulation methods and OFDM, radio localization methods, and radio resource management for ad-hoc and mobile networks.

Risto Wichman received his M.Sc. and D.Sc. (Tech) degrees in digital signal processing from Tampere University of Technology, Tampere, Finland, in 1990 and 1995, respectively.

From 1995 to 2001, he worked at Nokia Research Center as a senior research engineer. In 2002, he joined Department of Signal Processing and Acoustics, Faculty of Electrical Engineering, Aalto University, Finland, where he is a full professor since 2008. His research interests include signal processing techniques for wireless communication systems.

Publication 2

D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014. DOI: 10.1109/JSAC.2014.2330093

Widely Linear Digital Self-Interference Cancellation in Direct-Conversion Full-Duplex Transceiver

Dani Korpi, Lauri Anttila, Ville Syrjälä, and Mikko Valkama

***Abstract*—This article addresses the modeling and cancellation of self-interference in full-duplex direct-conversion radio transceivers, operating under practical imperfect radio frequency (RF) components. Firstly, detailed self-interference signal modeling is carried out, taking into account the most important RF imperfections, namely transmitter power amplifier nonlinear distortion as well as transmitter and receiver IQ mixer amplitude and phase imbalances. The analysis shows that after realistic antenna isolation and RF cancellation, the dominant self-interference waveform at receiver digital baseband can be modeled through a widely-linear transformation of the original transmit data, opposed to classical purely linear models. Such widely-linear self-interference waveform is physically stemming from the transmitter and receiver IQ imaging, and cannot be efficiently suppressed by classical linear digital cancellation. Motivated by this, novel widely-linear digital self-interference cancellation processing is then proposed and formulated, combined with efficient parameter estimation methods. Extensive simulation results demonstrate that the proposed widely-linear cancellation processing clearly outperforms the existing linear solutions, hence enabling the use of practical low-cost RF front-ends utilizing IQ mixing in full-duplex transceivers.**

***Index Terms*—Direct-conversion radio, full-duplex radio, self-interference, image frequency, IQ imbalance, widely-linear filtering**

I. Introduction

FULL-DUPLEX radio communications with simultaneous transmission and reception at the same radio frequency (RF) carrier is one of the emerging novel paradigms to improve the efficiency and flexibility of RF spectrum use. Some of the recent seminal works in this field are for example [1]–[8], to name a few. Practical realization and implementation of small and low-cost full-duplex transceivers, e.g., for mobile cellular radio or local area connectivity devices, are, however, still subject to many challenges. One of the biggest problems is the so called self-interference (SI), which is stemming from the simultaneous transmission and reception at single frequency, thus causing the strong transmit signal to couple directly to the receiver path.

In general, the transmitter and receiver may use either the same [9]–[11] or separate but closely-spaced antennas [2]–[5]. In this work, we focus on the case of separate antennas where depending on, e.g., the deployed physical antenna separation and transmit power, the level of the coupling SI signal can be in the order of 60–100 dB stronger than the actual received signal of interest at receiver input, especially when operating close to the sensitivity level of the receiver chain. To suppress such SI inside the transceiver, various antenna-based solutions [2], [4], [12], [13], active analog/RF cancellation methods [2]–[4], [8], [14]–[16] and digital baseband cancellation techniques [3], [14], [17], [18] have been proposed in the literature. In addition, a general analysis about the overall performance of different linear SI cancellation methods is performed in [7], while in [19], SI cancellation based on spatial-domain suppression is compared to subtractive time-domain cancellation, and rate regions are calculated for the two methods.

However, the performance of the SI cancellation mechanisms based on linear processing is usually limited by the analog/RF circuit non-idealities occurring within the full-duplex transceiver. For this reason, some of the most prominent types of such analog/RF circuit non-idealities have been analyzed in several recent studies. The phase noise of the transmitter and receiver oscillators has been analyzed in [20]–[22]. It was observed that the phase noise can significantly limit the amount of achievable SI suppression, especially when using two separate oscillators for transmitter and receiver. A signal model including the effect of phase noise is also investigated in [23], where the feasibility of asynchronous full-duplex communications is studied. In [24], the effects of the receiver chain noise figure and the quantization noise are also taken into account, in addition to phase noise. The authors then provide an approximation for the rate gain region of a full-duplex transceiver under the analyzed impairments. The effect of quantization noise is also analyzed in [25], where the relation between analog and digital cancellation under limited dynamic range for the analog-to-digital converter is studied. The existence of several non-idealities in the transmit chain, including power amplifier (PA) nonlinearity, is acknowledged in [26]. As a solution, the authors propose taking the cancellation signal from the output of the transmitter chain to exclude these non-idealities from the signal path. In [27], several non-idealities in the transmitter and receiver chains are also considered, including nonlinearities and IQ mismatch. Their effect is then studied in terms of the achievable SI suppression with linear processing and received spatial covariance eigenvalue distribution. In [28], the transmitter non-idealities are modelled as white noise, the power of which is dependent on the transmit power, and their effect is analyzed in the context of cognitive radios. In [29], comprehensive

D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, are with the Department of Electronics and Communications Engineering, Tampere University of Technology, PO Box 692, FI-33101, Tampere, Finland, e-mail: dani.korpi@tut.fi.

The research work leading to these results was funded by the Academy of Finland (under the project #259915 "In-band Full-Duplex MIMO Transmission: A Breakthrough to High-Speed Low-Latency Mobile Networks"), the Finnish Funding Agency for Technology and Innovation (Tekes, under the project "Full-Duplex Cognitive Radio"), the Linz Center of Mechatronics (LCM) in the framework of the Austrian COMET-K2 programme, and Emil Aaltonen Foundation.

distortion calculations, taking into account several sources of nonlinear distortion, among other things, are reported. The findings indicate that, especially with high transmit powers, the nonlinear distortion of the transmitter PA and receiver front-end components can significantly contribute to the SI waveform and thus hinder the efficiency of purely linear digital cancellation. Stemming from the findings of these studies, novel nonlinear cancellation processing solutions have been recently proposed in [14], [17], [18] to suppress such nonlinear SI, in addition to plain linear SI, at receiver digital baseband.

In addition to phase noise and nonlinearities, also other RF imperfections can impact the SI waveform and its cancellation. One particularly important imperfection in IQ processing based architectures is the so-called IQ imbalance and the corresponding inband IQ image or mirror component [30]. On the transmitter side, in general, such IQ imbalance is contributing to the transmitter error vector magnitude, and possibly also to adjacent channel leakage and spurious emissions, depending on the transmit architecture. As a practical example, 3GPP Long Term Evolution (LTE)/LTE-Advanced radio system specifications limit the minimum attenuation for the inband image component in mobile user equipment transmitters to 25 dB or 28 dB, depending on the specification release [31]. Such image attenuation is sufficient in the transmission path, but when considering the full-duplex device self-interference problem, the IQ image of the SI signal is additional interference leaking to the receiver path. Furthermore, additional IQ imaging of the SI signal takes place in the receiver path. In [32], the effect of IQ imbalance on a full-duplex transceiver was noticed in the measurements, as it was causing clear residual SI after all the cancellation stages. However, there is no previous work on compensating the SI mirror component caused by the IQ imbalance of the transmitter and receiver chains. Overall, there is relatively little discussion in the existing literature about the effect of IQ imbalance on the performance of a full-duplex transceiver. In part, this is due to the high-cost high-quality equipment that has been used to demonstrate the full-duplex transceiver principle in the existing implementations. For example, the WARP platform, which has been used at least in [2], [3], [5], [7], [14], provides an attenuation in the excess of 40–50 dB for the inband image component [33]. As our analysis in this article indicates, this is adequate to decrease the mirror component sufficiently low for it to have no significant effect on the performance of the full-duplex transceiver. Furthermore, if properly calibrated high-end laboratory equipment is used, the image attenuation can easily be even in the order of 60–80 dB, meaning that the effect of IQ imbalance is totally negligible. *However, for a typical mobile transceiver with low-cost mass-market analog/RF components, the image attenuation is generally significantly less, as already mentioned* [31]. This means that also the effect of IQ imbalance in a full-duplex transceiver must be analyzed and taken into account.

In this article, we firstly carry out detailed modeling for the SI waveform in different stages of the transmitter-receiver coupling path, taking into account the effects of transmitter and receiver IQ image components, as well as transmitter PA nonlinearities. Incorporating then also the effects of realistic multipath antenna coupling, linear analog/RF cancellation and linear digital baseband cancellation, the powers of the remaining SI components at the output of the full coupling and processing chain are analyzed. This analysis shows that, with realistic component values and linear cancellation processing, the IQ image of the classical linear SI is heavily limiting the receiver path signal-to-noise-plus-interference ratio (SINR). Such observation has not been reported earlier in the literature. Motived by these findings, a novel widely-linear (WL) digital SI canceller is then developed, where not only the original transmit data, but also its complex conjugate, modeling the IQ imaging, are processed to form an estimate of the SI signal. Efficient parameter estimation methods are also developed to estimate the cancellation parameters of the proposed WL structure through WL least-squares model fitting. The proposed WL SI canceller is shown by analysis, and through extensive simulations, to substantially improve the SI cancellation performance in the presence of practical IQ imaging levels, compared to classical purely linear processing, and it can hence enable full-duplex transceiver operation with realistically IQ balanced low-cost user equipment RF components.

The rest of this article is organized as follows. In Section II, the structure of the considered full-duplex transceiver and its baseband-equivalent model are presented, alongside with the overall self-interference signal model and its simplified version. Also principal system calculations, in terms of the powers of the different self-interference terms, are carried out. The proposed method for widely-linear digital cancellation and the estimation procedure for the coefficients are then presented in Section III. In Section IV, the performance of widely-linear digital cancellation under different scenarios is analyzed with full waveform simulations. Finally, the conclusions are drawn in Section V.

II. Full-Duplex Transceiver and Self-Interference

The structure of the analyzed full-duplex transceiver is presented in Fig. 1. It can be observed that the transceiver follows a typical direct-conversion architecture [34]–[36], which is well-known in previous literature, and thus it is not discussed here in detail. This architecture is chosen due to its simple structure and wide applications in modern wireless transceivers. The actual IQ imaging problem is caused by the IQ mixers at both the transmitter and receiver chains. Due to the inherent mismatches between the amplitudes and phases of the I- and Q-branches, the mirror image of the original signal is added on top of it, with certain image attenuation [30]. In this paper, we assume that the level of this image attenuation is similar to what is specified in 3GPP LTE specifications [31], namely 25 dB.

The actual analysis of the full-duplex transceiver and SI waveform at different stages of the transceiver is done next by using baseband-equivalent models. The block diagram of the overall baseband-equivalent model is shown in Fig. 2, alongside with the principal mathematical or behavioral model for each component, propagation of the transmitted signal, and

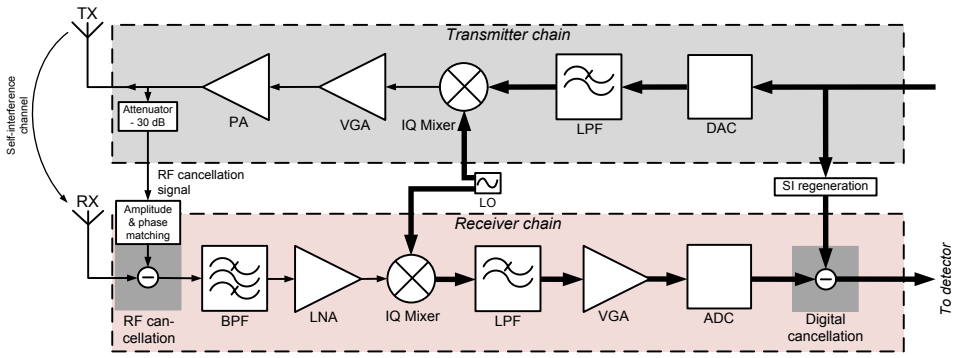


Fig. 1. A block diagram of the assumed direct-conversion full-duplex transceiver.

the corresponding variable names. In the following subsection, a complete characterization for the effective SI waveform in different stages of the transceiver is provided, using the same notations as in Fig. 2, taking into account transmitter IQ imaging and PA distortion, realistic multipath coupling channel, realistic linear analog/RF cancellation, receiver IQ imaging, and receiver linear cancellation. Stemming from this, the powers of the different SI terms at receiver digital baseband are then analysed in Subsection II-B.

A. Self-Interference Signal Model with Practical RF Components

The complex baseband transmitted signal is denoted by $x(n)$, or by $x(t)$ after digital-to-analog conversion. It is assumed that the power of $x(t)$ is such that the desired transmit power is reached after the amplification by IQ mixer and PA, i.e., the transmitter VGA is omitted from the baseband-equivalent model. This is done to make the notation simpler and thus more illustrative. In addition, the signal $x(n)$ is perfectly known, as it is generated within the transceiver. In this analysis, the digital-to-analog converters (DACs) and low-pass filters (LPFs) are assumed to be ideal, but the IQ mixer is assumed to have some imbalance between the I and Q branches. The signal at the output of the transmit IQ mixer can be now written as

$$x_{IQ}^{TX}(t) = g_{1,TX}(t) \star x(t) + g_{2,TX}(t) \star x^*(t), \quad (1)$$

where $g_{1,TX}(t)$ is the response for the direct signal component, and $g_{2,TX}(t)$ is the response for the image component [30]. Here $(\cdot)^*$ indicates the complex conjugate and $\star$ denotes the convolution operation. Above-kind of transformations, where both direct and complex-conjugated signals are filtered and finally summed together, are typically called widely-linear in the literature, see e.g. [37], [38].

The quality of the IQ mixer can be quantified by the image rejection ratio (IRR). With the variables used in this analysis, it can be defined for the transmitter as

$$IRR_{TX}(f) = \frac{|G_{1,TX}(f)|^2}{|G_{2,TX}(f)|^2}, \quad (2)$$

where $G_{1,TX}(f)$ and $G_{2,TX}(f)$ are the frequency-domain representations of $g_{1,TX}(t)$ and $g_{2,TX}(t)$, respectively [30]. A similar characterization can obviously be established also for the receiver IQ mixer, referred to as $IRR_{RX}(f)$.

Before transmission, the signal is amplified with a nonlinear PA. In this analysis, we model the PA response with a Hammerstein nonlinearity [39], [40] given as

$$x_{PA}(t) = \left(\alpha_0 x_{IQ}^{TX}(t) + \alpha_1 x_{IQ}^{TX}(t)|x_{IQ}^{TX}(t)|^2\right) \star f(t), \quad (3)$$

where α_0 is the linear gain, α_1 is the gain of the third order component, and $f(t)$ is the memory model of the PA. For simplicity, we write $x_{IMD}(t) = x_{IQ}^{TX}(t)|x_{IQ}^{TX}(t)|^2$, and use this to refer to the third order nonlinear component. Thus, we can rewrite (3) as

$$x_{PA}(t) = \left(\alpha_0 x_{IQ}^{TX}(t) + \alpha_1 x_{IMD}(t)\right) \star f(t). \quad (4)$$

It is obvious that true PAs contain also distortion components beyond third-order, but in this analysis we make a simplification and focus only on the third-order distortion, as that is in practice always the strongest nonlinearity at PA output.

The transmitted signal is next coupled back to the receive antenna, thus producing SI. In this analysis, to simplify the notations, it is assumed that there is no actual received signal of interest. This will decrease the complexity of the equations while having no significant effect on the results, as the purpose is to characterize the final SI waveform at receiver baseband. Thus, the signal at the input of the receiver chain is of the following form:

$$y(t) = h_{ch}(t) \star x_{PA}(t) + n_{th}(t), \quad (5)$$

where $h_{ch}(t)$ is the multipath coupling channel between transmit and receive antennas, and $n_{th}(t)$ denotes thermal noise. To suppress the SI before it enters the LNA, RF cancellation is performed. The signal after RF cancellation can be expressed as

$$y_{RF}(t) = y(t) - a(t) \star x_{PA}(t), \quad (6)$$

where $a(t)$ is typically an estimate for the main path of the coupling channel [2], [3], [7]. In other words, $a(t)$ is a one tap filter, depicting the delay, phase, and attenuation of the main

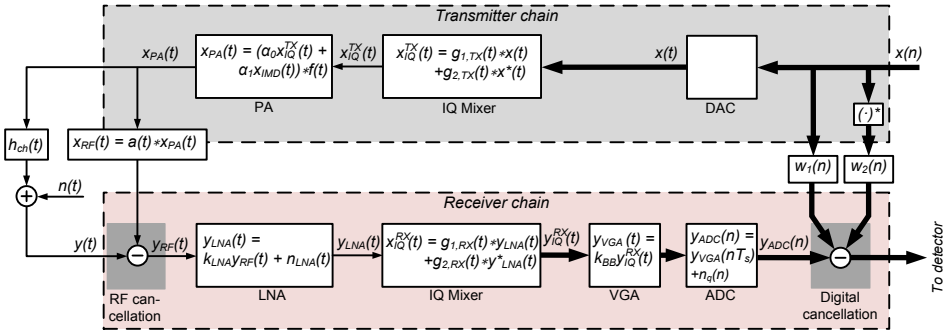


Fig. 2. A baseband-equivalent model of the analyzed full-duplex transceiver, alongside with the signals propagating at the different stages of the transceiver. Digital baseband SI cancellation is already showing, at structural level, the proposed WL SI cancellation where both the original transmit data and its complex conjugate are processed.

coupling propagation. Thus, in this analysis it is assumed that RF cancellation attenuates only the direct coupling component. Furthermore, in the numerical experiments and results, $a(t)$ is chosen so that it provides the desired amount of SI attenuation at RF, e.g., 30 dB, as has been reported in practical experiments [5], [7], [15]. This is done by tuning the error of its attenuation and delay, modeling realistic RF cancellation.

Next, the received signal is amplified by the low-noise amplifier (LNA). The output signal of the LNA can be written as

$$y_{LNA}(t) = k_{LNA} y_{RF}(t) + n_{LNA}(t), \tag{7}$$

where k_{LNA} is the complex gain of the LNA, and $n_{LNA}(t)$ is the increase in the noise floor caused by the LNA. The nonlinearity effects of the receiver chain amplifiers are not taken into account in this paper, as they are insignificant in comparison to the other distortion components under typical circumstances [29]. This will also significantly simplify the analysis.

Similar to the transmitter, the receiver IQ mixer has IQ imbalance, and thus it produces an image component of the total signal entering the mixer, including the SI. The signal at the output of the receiver IQ mixer can be now expressed as

$$y_{IQ}^{RX}(t) = g_{1,RX}(t) \star y_{LNA}(t) + g_{2,RX}(t) \star y^*_{LNA}(t), \tag{8}$$

where $g_{1,RX}(t)$ is the response for the direct signal component, and $g_{2,RX}(t)$ is the response for the receiver image component.

Finally, the signal is amplified by the variable gain amplifier (VGA) to match its waveform dynamics to the voltage range of the analog-to-digital converter (ADC) and then digitized. The digitized signal can be written as

$$y_{ADC}(nT_s) = k_{BB} y_{IQ}^{RX}(nT_s) + n_q(nT_s), \tag{9}$$

where k_{BB} is the complex baseband gain of the VGA, T_s is the sampling time, and $n_q(nT_s)$ denotes quantization noise. In the continuation, we drop the sampling interval T_s from the equations for brevity and use only the discrete-time index n.

To express the residual SI signal at the digital domain in terms of the known transmit data $x(n)$, a complete equation for $y_{ADC}(n)$ is next derived by substituting (1) to (4), (4) to (5) and so on. After these elementary manipulations, we arrive at the following equation for $y_{ADC}(n)$, with respect to $x(n)$ and fundamental system responses:

$$\begin{aligned} y_{ADC}(n) = {} & h_1(n) \star x(n) + h_2(n) \star x^*(n) \\ & + h_{IMD}(n) \star x_{IMD}(n) + h_{IMD,im}(n) \star x^*_{IMD}(n) \\ & + h_n(n) \star n_{tot}(n) + h_{n,im}(n) \star n^*_{tot}(n) + n_q(n), \end{aligned} \tag{10}$$

where we have defined a total noise signal as $n_{tot}(n) = k_{LNA} n_{th}(n) + n_{LNA}(n)$, including the thermal noise at the input of the receiver chain and the additional noise produced by the LNA. The channel responses of the individual signal components can be written as follows:

$$\begin{aligned} h_1(n) = {} & k_{BB} k_{LNA} \alpha_0 g_{1,TX}(n) \star g_{1,RX}(n) \star f(n) \\ & \star (h_{ch}(n) - a(n)) + k_{BB} k^*_{LNA} \alpha^*_0 g^*_{2,TX}(n) \\ & \star g_{2,RX}(n) \star f^*(n) \star (h^*_{ch}(n) - a^*(n)) \end{aligned} \tag{11}$$

$$\begin{aligned} h_2(n) = {} & k_{BB} k_{LNA} \alpha_0 g_{2,TX}(n) \star g_{1,RX}(n) \star f(n) \\ & \star (h_{ch}(n) - a(n)) + k_{BB} k^*_{LNA} \alpha^*_0 g^*_{1,TX}(n) \\ & \star g_{2,RX}(n) \star f^*(n) \star (h^*_{ch}(n) - a^*(n)) \end{aligned} \tag{12}$$

$$\begin{aligned} h_{IMD}(n) = {} & k_{BB} k_{LNA} \alpha_1 g_{1,RX}(n) \star f(n) \\ & \star (h_{ch}(n) - a(n)) \end{aligned} \tag{13}$$

$$\begin{aligned} h_{IMD,im}(n) = {} & k_{BB} k^*_{LNA} \alpha^*_1 g_{2,RX}(n) \star f^*(n) \\ & \star (h^*_{ch}(n) - a^*(n)) \end{aligned} \tag{14}$$

$$h_n(n) = k_{BB} g_{1,RX}(n) \tag{15}$$

$$h_{n,im}(n) = k_{BB} g_{2,RX}(n). \tag{16}$$

As can be seen in (10), the total SI at receiver digital baseband contains not only the linear SI but also its complex conjugate. These different components of the SI signal are hereinafter referred to as linear SI and conjugate SI, respectively. In addition to these signal components, PA-induced IMD and its

complex-conjugate, which will similarly be referred to as IMD and conjugate IMD, are also present in the total SI signal.

Using the above equations, it is possible to describe the effect of conventional linear digital SI cancellation, which can attenuate only the linear SI component. Corresponding to the notation in Fig. 2, where the linear channel estimate is denoted by $w_1(n)$, the signal after linear digital cancellation can be expressed as

$$\begin{aligned} y_{LDC}(n) &= y_{ADC}(n) - w_1(n) \star x(n) \\ &= (h_1(n) - w_1(n)) \star x(n) + h_2(n) \star x^*(n) \\ &+ h_{IMD}(n) \star x_{IMD}(n) + h_{IMD,im}(n) \star x^*_{IMD}(n) \\ &+ h_n(n) \star n_{tot}(n) + h_{n,im}(n) \star n^*_{tot}(n) + n_q(n). \end{aligned} \tag{17}$$

From (17) it can be observed that $w_1(n)$ should estimate the channel of the linear SI component, i.e., $w_1(n) = \hat{h}_1(n)$. However, even with a perfect estimate of $h_1(n)$, the signal $y_{LDC}(n)$ can still be substantial interference from weak desired signal perspective. This is mainly due to the conjugate SI, IMD, and conjugate IMD. Obviously, there is also some thermal noise, but typically it is not significantly limiting the performance of a full-duplex transceiver.

Notice that nonlinear distortion has been shown earlier in the literature to limit the achievable SINR of a full-duplex radio, and there are also methods for attenuating it [14], [17], [18]. However, there is no previous work on analyzing and attenuating the conjugate SI signal, which is relative to $x^*(n)$ in our notations. In the next subsection, we will analyze the relative strength of this conjugate SI through principal power calculations, and show that with typical RF component specifications, it is the dominant SINR limiting phenomenon. In Section III, we then also provide a method for suppressing the conjugate SI in the digital domain by processing the original transmit data in a widely-linear manner with two filters $w_1(n)$ and $w_2(n)$, marked also in Fig. 2. However, in the following subsection it is still first assumed that $w_2(n) = 0$, and thus no compensation is done for the conjugate SI, in order to properly quantify and illustrate the limitations of classical linear SI cancellation.

B. Principal System Calculations for Different Distortion Terms

In order to analyze and illustrate the relative levels of the different distortion components of the overall SI signal, a somewhat simplified scenario is first considered. More specifically, the frequency-dependent characteristics of different distortion components are neglected, which allows for the equations to be presented in a more compact and illustrative form. Furthermore, as we are here primarily interested in the average powers of different distortion components, neglecting inband frequency-dependency is well justified. If we now denote the general impulse function by $\delta(n)$, the different system impulse responses of the distortion components can be expressed as follows: $g_{1,TX}(n) \approx g_{1,TX}\delta(n)$, $g_{2,TX}(n) \approx g_{2,TX}\delta(n)$, $g_{1,RX}(n) \approx g_{1,RX}\delta(n)$, $g_{2,RX}(n) \approx g_{2,RX}\delta(n)$, $f(n) \approx \delta(n)$, $h_{ch}(n) \approx h_{ch}\delta(n)$, $a(n) \approx a\delta(n)$, and $w_1(n) \approx w_1\delta(n)$. By substituting these simplified terms into (11)–(16), we can rewrite them as

$$\begin{aligned} h_1(n) &\approx k_{BB}k_{LNA}\alpha_0 g_{1,TX}g_{1,RX}(h_{ch} - a)\delta(n) \\ &+ k_{BB}k^*_{LNA}\alpha^*_0 g^*_{2,TX}g_{2,RX}(h^*_{ch} - a^*)\delta(n) \end{aligned} \tag{18}$$

$$\begin{aligned} h_2(n) &\approx k_{BB}k_{LNA}\alpha_0 g_{2,TX}g_{1,RX}(h_{ch} - a)\delta(n) \\ &+ k_{BB}k^*_{LNA}\alpha^*_0 g^*_{1,TX}g_{2,RX}(h^*_{ch} - a^*)\delta(n) \end{aligned} \tag{19}$$

$$h_{IMD}(n) \approx k_{BB}k_{LNA}\alpha_1 g_{1,RX}(h_{ch} - a)\delta(n) \tag{20}$$

$$h_{IMD,im}(n) \approx k_{BB}k^*_{LNA}\alpha^*_1 g_{2,RX}(h^*_{ch} - a^*)\delta(n) \tag{21}$$

$$h_n(n) \approx k_{BB}g_{1,RX}\delta(n) \tag{22}$$

$$h_{n,im}(n) \approx k_{BB}g_{2,RX}\delta(n). \tag{23}$$

Furthermore, as the magnitude of the term $g^*_{2,TX}g_{2,RX}$ is very small in comparison to the magnitude of $g_{1,TX}g_{1,RX}$, even when considering a relatively low image attenuation, we can write (18) as

$$h_1(n) \approx k_{BB}k_{LNA}\alpha_0 g_{1,TX}g_{1,RX}(h_{ch} - a)\delta(n) \tag{24}$$

Using (19)–(24), $y_{LDC}(n)$ can then be expressed in a simplified form as follows:

$$\begin{aligned} y_{LDC}(n) &\approx (h_1 - w_1)x(n) + h_2x^*(n) + h_{IMD}x_{IMD}(n) \\ &+ h_{IMD,im}x^*_{IMD}(n) + h_n n_{tot}(n) + h_{n,im}n^*_{tot}(n) \\ &+ n_q(n) \\ &= (k_{BB}k_{LNA}\alpha_0 g_{1,TX}g_{1,RX}(h_{ch} - a) - w_1)x(n) \\ &+ (k_{BB}k_{LNA}\alpha_0 g_{2,TX}g_{1,RX}(h_{ch} - a) \\ &+ k_{BB}k^*_{LNA}\alpha^*_0 g^*_{1,TX}g_{2,RX}(h^*_{ch} - a^*))x^*(n) \\ &+ k_{BB}k_{LNA}\alpha_1 g_{1,RX}(h_{ch} - a)x_{IMD}(n) \\ &+ k_{BB}k^*_{LNA}\alpha^*_1 g_{2,RX}(h^*_{ch} - a^*)x^*_{IMD}(n) \\ &+ k_{BB}g_{1,RX}n_{tot}(n) + k_{BB}g_{2,RX}n^*_{tot}(n) + n_q(n). \end{aligned} \tag{25}$$

From (25), it is now possible to calculate the powers of the different signal components. These powers are defined as follows:

$$\begin{aligned} p_{SI} &= E[|k_{BB}k_{LNA}\alpha_0 g_{1,TX}g_{1,RX}(h_{ch} - a) - w_1|^2|x(n)|^2] \\ &= E[|k_{BB}k_{LNA}\alpha_0 g_{1,TX}g_{1,RX}(h_{ch} - a) - w_1|^2] \\ &\times E[|x(n)|^2] \\ &= E[|k_{BB}k_{LNA}\alpha_0 g_{1,TX}g_{1,RX}(h_{ch} - a) - w_1|^2]p_x \end{aligned} \tag{26}$$

$$\begin{aligned} p_{SI,im} &= E[|k_{BB}k_{LNA}\alpha_0 g_{2,TX}g_{1,RX}(h_{ch} - a) \\ &+ k_{BB}k^*_{LNA}\alpha^*_0 g^*_{1,TX}g_{2,RX}(h^*_{ch} - a^*)|^2]p_x \end{aligned} \tag{27}$$

$$p_{IMD} = E[|k_{BB}k_{LNA}\alpha_1 g_{1,RX}(h_{ch} - a)|^2]p_{x,IMD} \tag{28}$$

$$p_{IMD,im} = E[|k_{BB}k^*_{LNA}\alpha^*_1 g_{2,RX}(h^*_{ch} - a^*)|^2]p_{x,IMD} \tag{29}$$

$$p_{noise} = E[|k_{BB}g_{1,RX}|^2]p_n \tag{30}$$

$$p_{noise,im} = E[|k_{BB}g_{2,RX}|^2]p_n \tag{31}$$

$$p_q = E[|n_q(n)|^2] = \frac{p_{AD}}{snr_{ADC}}, \tag{32}$$

where $E[\cdot]$ denotes the expected value, $p_x = E[|x(n)|^2]$ is the power of the signals $x(n)$ and $x^*(n)$, as conjugation does

not affect the power of the signal, $p_{x,IMD} = E[|x_{IMD}(n)|^2]$ is the power of the signals $x_{IMD}(n)$ and $x^*_{IMD}(n)$, $p_n = E[|n_{tot}(n)|^2]$ is the power of the signals $n_{tot}(n)$ and $n^*_{tot}(n)$, and $p_q = E[|n_q(n)|^2]$ is the power of quantization noise. Here, p_q is defined in terms of the total signal power at the input of the ADC, p_{AD}, and the SNR of the ADC, $snr_{ADC} = 10^{(6.02b+4.76-PAPR)/10}$, where b is the number of bits at the ADC, and $PAPR$ is the peak-to-average-power ratio in dB [34].

For further simplicity, the power of the total noise term $p_n = E[|n_{tot}(n)|^2]$ can be expressed in a more compact manner. As the LNA constitutes for most of the noise factor of the receiver chain, denoted by F, it can be appoximated with very little error that LNA has a noise factor of F, and no additional noise is produced throughout the rest of the receiver chain. Thus, if the power of the thermal noise at the input of the receiver chain is defined as $p_{th} = E[|n_{th}(n)|^2]$, we can write $p_n = E[|n_{tot}(n)|^2] = E[|k_{LNA}n_{th}(n) + n_{LNA}(n)|^2] = |k_{LNA}|^2 F p_{th}$, based on the definition of the noise factor [34].

In addition, to express the power levels using the defined parameters, the amount of achieved RF cancellation with respect to the error between h_{ch} and a must be defined. This can be done by comparing the power of the total SI signal before and after RF cancellation. Based on (5), the SI signal before RF cancellation is $y_{SI}(t) = h_{ch}(t) \star x_{PA}(t)$. After RF cancellation, the SI signal is $y_{RF,SI}(t) = y_{SI}(t) - a(t) \star x_{PA}(t) = (h_{ch}(t) - a(t)) \star x_{PA}(t)$, based on (6). Taking the previously defined simplifications into account, we can write $y_{SI}(t) \approx h_{ch}x_{PA}(t)$ and $y_{RF,SI}(t) = (h_{ch} - a)x_{PA}(t)$. Thus, the amount of RF cancellation can be defined as

$$|a_{RF}|^2 = \frac{E[|y_{RF,SI}(t)|^2]}{E[|y_{SI}(t)|^2]} = \frac{E[|(h_{ch}-a)x_{PA}(t)|^2]}{E[|h_{ch}x_{PA}(t)|^2]} = \frac{E[|h_{ch}-a|^2]E[|x_{PA}(t)|^2]}{E[|h_{ch}|^2]E[|x_{PA}(t)|^2]} = \frac{E[|h_{ch}-a|^2]}{E[|h_{ch}|^2]}. \quad (33)$$

Here it is assumed that the instantaneous path loss of the SI coupling channel, denoted by $|h_{ch}|^2$, and the error of the RF cancellation channel estimate, denoted by $(h_{ch} - a)$, are circular and normally distributed. The amount of antenna attenuation is now defined as $|a_{ant}|^2 = E[|h_{ch}|^2]$. Using (33), it is then possible to define the power of the error between the SI coupling channel and the channel estimate for RF cancellation as

$$E[|h_{ch} - a|^2] = E[|h_{ch}|^2]|a_{RF}|^2 = |a_{ant}|^2|a_{RF}|^2. \quad (34)$$

The amount of achieved linear digital cancellation is defined next as the decrease in the power of the linear SI component $x(n)$. Before digital cancellation, the linear SI signal can be expressed as $y_{ADC,SI}(n) = h_1(n) \star x(n) \approx h_1 x(n)$, and after linear digital cancellation as $y_{LDC,SI}(n) = y_{ADC,SI}(n) - w_1(n) \star x(n) \approx (h_1 - w_1)x(n)$. The attenuation of the linear SI power by linear digital cancellation can then be expressed as follows:

$$|a_{LDC}|^2 = \frac{E[|y_{LDC,SI}(n)|^2]}{E[|y_{ADC,SI}(n)|^2]} = \frac{E[|(h_1-w_1)x(n)|^2]}{E[|h_1x(n)|^2]} = \frac{E[|h_1-w_1|^2]E[|x(n)|^2]}{E[|h_1|^2]E[|x(n)|^2]} = \frac{E[|h_1-w_1|^2]}{E[|h_1|^2]}. \quad (35)$$

Now, by using (35), it is possible to express $E[|h_1 - w_1|^2]$ in terms of $|a_{LDC}|^2$ as follows:

$$E[|h_1 - w_1|^2] = |a_{LDC}|^2 E[|h_1|^2], \quad (36)$$

where $h_1 = k_{BB}k_{LNA}\alpha_0 g_{1,TX}g_{1,RX}(h_{ch} - a)$.

Now, by substituting (34) to (26)–(29), (36) to (26), and $p_n = |k_{LNA}|^2 F p_{th}$ to (30)–(31), we can finally express the power levels of all the different signal components as follows:

$$\begin{aligned} p_{SI} &= |a_{LDC}|^2 E[|k_{BB}|^2|k_{LNA}|^2|\alpha_0|^2|h_{ch}-a|^2 \\ &\quad \times |g_{1,TX}|^2|g_{1,RX}|^2]p_x \\ &= |a_{LDC}|^2|k_{BB}|^2|k_{LNA}|^2|\alpha_0|^2 E[|h_{ch}-a|^2] \\ &\quad \times |g_{1,TX}|^2|g_{1,RX}|^2 p_x \\ &= |a_{LDC}|^2|k_{BB}|^2|k_{LNA}|^2|\alpha_0|^2|a_{ant}|^2|a_{RF}|^2 \\ &\quad \times |g_{1,TX}|^2|g_{1,RX}|^2 p_x \end{aligned} \quad (37)$$

$$\begin{aligned} p_{SI,im} &= (E[|k_{BB}|^2|k_{LNA}|^2|\alpha_0|^2|g_{2,TX}|^2|g_{1,RX}|^2 \\ &\quad \times |h_{ch}-a|^2] + E[|k_{BB}|^2|k^*_{LNA}|^2|\alpha^*_0|^2|g^*_{1,TX}|^2 \\ &\quad \times |g_{2,RX}|^2|h^*_{ch}-a^*|^2] + 2E[Re\{|k_{BB}|^2 k^2_{LNA}\alpha^2_0 \\ &\quad \times g_{1,TX}g_{2,TX}g_{1,RX}g^*_{2,RX}(h_{ch}-a)^2\}])p_x \\ &= (|k_{BB}|^2|k_{LNA}|^2|\alpha_0|^2|g_{2,TX}|^2|g_{1,RX}|^2 \\ &\quad \times E[|h_{ch}-a|^2] + |k_{BB}|^2|k^*_{LNA}|^2|\alpha^*_0|^2|g^*_{1,TX}|^2 \\ &\quad \times |g_{2,RX}|^2 E[|h^*_{ch}-a^*|^2] + 2Re\{|k_{BB}|^2 k^2_{LNA}\alpha^2_0 \\ &\quad \times g_{1,TX}g_{2,TX}g_{1,RX}g^*_{2,RX}E[(h_{ch}-a)^2]\})p_x \\ &= (|k_{BB}|^2|k_{LNA}|^2|\alpha_0|^2|a_{ant}|^2|a_{RF}|^2 \\ &\quad \times (|g_{2,TX}|^2|g_{1,RX}|^2 + |g_{1,TX}|^2|g_{2,RX}|^2)p_x \end{aligned} \quad (38)$$

$$\begin{aligned} p_{IMD} &= |k_{BB}|^2|k_{LNA}|^2|\alpha_1|^2|g_{1,RX}|^2 \\ &\quad \times |a_{ant}|^2|a_{RF}|^2 p_{x,IMD} \end{aligned} \quad (39)$$

$$\begin{aligned} p_{IMD,im} &= |k_{BB}|^2|k_{LNA}|^2|\alpha_1|^2|g_{2,RX}|^2 \\ &\quad \times |a_{ant}|^2|a_{RF}|^2 p_{x,IMD} \end{aligned} \quad (40)$$

$$p_{noise} = F|k_{BB}|^2|k_{LNA}|^2|g_{1,RX}|^2 p_{th} \quad (41)$$

$$p_{noise,im} = F|k_{BB}|^2|k_{LNA}|^2|g_{2,RX}|^2 p_{th} \quad (42)$$

$$p_q = \frac{p_{AD}}{snr_{ADC}}. \quad (43)$$

In the above equations it is assumed that the complex gains of the different RF components are static and deterministic, whereas the error of the RF channel estimate is assumed to be a circular random variable, as explained earlier. Furthermore, the final term of the first equation for $p_{SI,im}$ can be omitted, as $E[(h_{ch} - a)^2] = 0$ due to the circularity assumption. The above set of derived formulas in (37)–(43) can now be used to evaluate the powers of different distortion terms at detector input, and in particular how they depend on the transmit power, antenna isolation, active RF cancellation and linear digital cancellation as well as on the transceiver RF imperfections.

System Calculations Example: To illustrate the relative strengths of the different signal components, typical component parameters are chosen, and (37)–(43) are used to determine the corresponding power levels, which will then be shown for a specified transmit power range. The used parameters are listed in Table I, and they correspond to a typical

TABLE I
EXAMPLE SYSTEM LEVEL PARAMETERS FOR THE FULL-DUPLEX TRANSCEIVER.

Parameter	Value
Bandwidth	12.5 MHz
Thermal noise floor at receiver input	-103.0 dBm
Receiver noise figure	4.1 dB
SNR requirement at detector input	10 dB
Sensitivity level	-88.9 dBm
Power of the received signal	-83.9 dBm
Transmit power	varied
PA gain	27 dB
PA IIP3	20 dBm
Antenna separation	40 dB
RF cancellation	30 dB
LNA gain	25 dB
IQ mixer gain (RX and TX)	6 dB
IRR (RX and TX)	25 dB
VGA gain (RX)	1–51 dB
ADC bits	12
ADC P-P voltage range	4.5 V
PAPR	10 dB

wideband transceiver with low-cost mass-product components. The value for IRR, describing the image attenuation, is chosen based on 3GPP LTE specifications [31]. The baseband VGA of the receiver chain is assumed to match the total waveform dynamics at the ADC input to the available voltage range.

In addition to the distortion powers, the power of the received signal of interest at detector input, denoted by p_{SOI}, is also shown in the figures as a reference, although it is not included in the signal model. This is done in order to be able to put the various distortion powers to proper context. The power of the signal of interest is chosen to be 5 dB above the sensitivity level at the input of the receiver chain. As the receiver sensitivity is defined for a 10 dB thermal noise SNR at detector input, this means that the signal of interest will be 15 dB above the thermal noise floor in the digital domain, which is then also the detector input SINR if no SI is present.

The power levels of the different signal components after linear digital cancellation, for transmit powers from -5 dBm to 25 dBm, are shown in Fig. 3. The amount of digital cancellation is chosen so that linear SI is attenuated below the thermal noise floor. This has been observed to be close to the true performance of digital cancellation under realistic conditions [14], [29]. In this example, this requires 27–57 dB of linear digital cancellation, depending on the transmit power.

From Fig. 3 it can be observed that the conjugate SI signal is clearly the most dominant distortion under a wide range of transmit powers. Actually, with transmit powers above 9 dBm, the power of the conjugate SI is even more powerful than the power of the signal of interest. Thus, with the chosen parameters, the conjugate SI is seriously degrading the achievable SINR of the full-duplex transceiver, which motivates the study of possible methods for attenuating it. Now, the ideal SINR of 15 dB is unreachable with the whole transmit power range from -5 dBm onwards, due to the powerful conjugate SI.

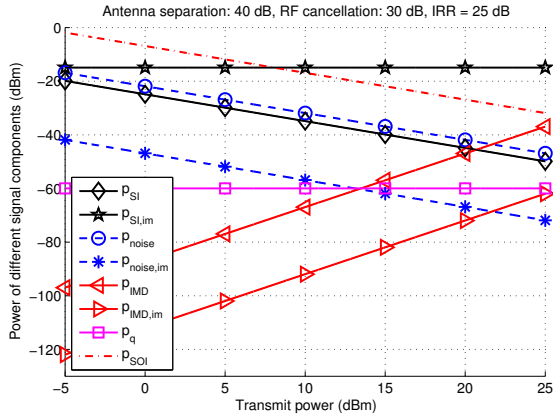


Fig. 3. The power levels of the different signal components after linear digital cancellation.

To investigate a different scenario, the amount of antenna separation is next assumed to be 30 dB, and the amount of RF cancellation 20 dB, which are achievable figures, even with very small antenna distance and low quality components for RF cancellation [2], [4]. In addition, the IRR of the IQ mixers is increased to 35 dB to model the effect of using expensive, higher quality IQ mixers. Again, the amount of digital cancellation is chosen so that it attenuates the linear SI below the thermal noise, now requiring 47–77 dB of linear digital SI attenuation.

The resulting power levels are shown in Fig. 4. Now it can be observed that the power of the conjugate SI is even higher with respect to the power of the signal of interest, due to less analog SI cancellation. Thus, with more pessimistic values for antenna separation and RF cancellation, the effect of conjugate SI is very severe, even when using higher quality IQ mixers. It should be noted, however, that with transmit powers above 15 dBm, also the IMD produced by the PA can be observed to be a significant factor. This indicates that, in order to achieve higher transmit powers with these parameters, attenuating only the linear and conjugate SI may not be sufficient, as also the nonlinear distortion will degrade the signal quality.

III. PROPOSED WIDELY-LINEAR DIGITAL CANCELLATION

It was observed in the previous section that, with typical component parameters, conjugate SI is the dominating source of distortion after classical linear digital cancellation. Thus, the performance of the analyzed full-duplex transceiver can be enhanced by attenuating also the SI image component in the cancellation processing. This can be done by utilizing widely-linear digital SI cancellation, the principle of which is proposed and formulated in this section.

In order to focus on the IQ image induced problematics, we assume below that the IMD produced by the PA, alongside with the conjugate IMD, are clearly weaker than the conjugate SI. This is also clearly visible in Fig. 3. We wish to acknowledge, though, that under certain circumstances, also the PA-induced nonlinear SI component must be further suppressed

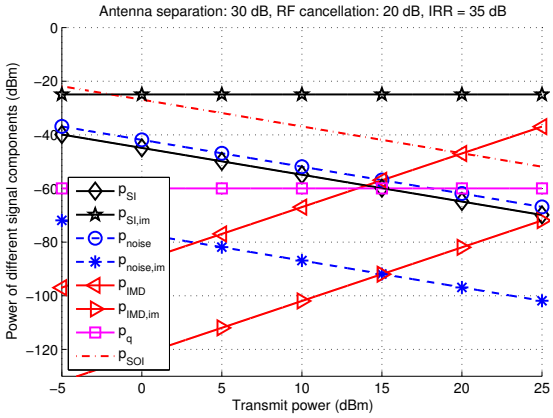


Fig. 4. The power levels of the different signal components after linear digital cancellation with the altered parameters.

[14], [17], [18], but below we focus for simplicity on the dominating conjugate SI stemming from the IQ imbalances. A combined canceller for suppressing both the IMD and the conjugate SI is left for future work.

A. Widely-Linear Cancellation Principle

Our starting point is the fundamental signal model in (10), which we can first re-write as

$$y_{ADC}(n) = h_1(n) \star x(n) + h_2(n) \star x^*(n) + z(n),$$

where

$$\begin{aligned} h_1(n) &= k_{BB} k_{LNA} \alpha_0 g_{1,TX}(n) \star g_{1,RX}(n) \\ &\quad \star (h_{ch}(n) - a(n)) + k_{BB} k^*_{LNA} \alpha^*_0 g^*_{2,TX}(n) \\ &\quad \star g_{2,RX}(n) \star (h^*_{ch}(n) - a^*(n)), \end{aligned}$$

$$\begin{aligned} h_2(n) &= k_{BB} k_{LNA} \alpha_0 g_{2,TX}(n) \star g_{1,RX}(n) \\ &\quad \star (h_{ch}(n) - a(n)) + k_{BB} k^*_{LNA} \alpha^*_0 g^*_{1,TX}(n) \\ &\quad \star g_{2,RX}(n) \star (h^*_{ch}(n) - a^*(n)), \end{aligned}$$

and $z(n)$ denotes in general the sum of all other signal terms, most notably thermal noise and quantization noise, and also PA-induced nonlinear distortion if present. Notice that here the general case of frequency-dependent antenna coupling and frequency-dependent IQ imbalances is considered, without any approximations.

The cancellation of the conjugate SI can be done in a similar manner as the cancellation of the direct component in (17). In this case, however, the known transmitted signal must first be conjugated, and the conjugated samples are then filtered with the channel estimate of the effective image channel, $h_2(n)$, to produce the cancellation signal. If the estimate for the channel of the conjugate SI signal is denoted by $w_2(n)$, corresponding to the notation in Fig. 2, then the total signal after the cancellation stage can be expressed as

$$\begin{aligned} y_{WLDC}(n) &= y_{ADC}(n) - w_1(n) \star x(n) - w_2(n) \star x^*(n) \\ &= h_1(n) \star x(n) - w_1(n) \star x(n) + h_2(n) \star x^*(n) \\ &\quad - w_2(n) \star x^*(n) + z(n) \\ &= (h_1(n) - w_1(n)) \star x(n) \\ &\quad + (h_2(n) - w_2(n)) \star x^*(n) + z(n) \end{aligned} \tag{44}$$

Thus, by estimating $h_2(n)$, and convolving the conjugated transmit samples with it, a cancellation signal for the conjugate SI can be obtained. This type of cancellation procedure, which takes into account also the mirror image component, in addition to the direct linear component, is referred to as widely-linear (WL) digital cancellation as it is processing both the direct transmit data as well as its complex conjugate [37], [38].

B. Widely-Linear Least-Squares Parameter Estimation

The parameters required for the proposed WL digital cancellation can be obtained with ordinary least-squares estimation where the samples of the observed signal $y_{ADC}(n)$ serve as the reference and two estimation filters are fitted to it through the known transmit data $x(n)$ and its complex conjugate $x^*(n)$. In the following, to allow a more articulate mathematical description of the parameter estimation and the corresponding cancellation procedure, vector-matrix notations are used.

At digital baseband, the observed signal is $y_{ADC}(n) = h_1(n) \star x(n) + h_2(n) \star x^*(n) + z(n)$. Stacking the signals at hand into column vectors over an observation period of N samples, this can be written as $\mathbf{y}_{ADC} = \mathbf{H}_1\mathbf{x} + \mathbf{H}_2\mathbf{x}^* + \mathbf{z}$, where $\mathbf{H}_1$ and $\mathbf{H}_2$ denote convolution matrices. Due to the commutability of the convolution operation, $\mathbf{y}_{ADC}$ can also be expressed as

$$\begin{aligned} \mathbf{y}_{ADC} &= \mathbf{X}\mathbf{h}_1 + \mathbf{X}^*\mathbf{h}_2 + \mathbf{z} = \begin{bmatrix} \mathbf{X} & \mathbf{X}^* \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix} + \mathbf{z} \\ &= \mathbf{X}_{aug}\mathbf{h}_{aug} + \mathbf{z}, \end{aligned} \tag{45}$$

where $\mathbf{y}_{ADC} = [y_{ADC}(M-1)\ y_{ADC}(M) \cdots y_{ADC}(N-K-1)]^T$, and the covariance windowed convolution data matrix $\mathbf{X}$, with K upper rows removed, is of the form

$$\mathbf{X} = \begin{bmatrix} x(M+K-1) & x(M+K-2) & \cdots & x(K) \\ x(M+K) & x(M+K-1) & \cdots & x(K+1) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-1) & x(N-2) & \cdots & x(N-M) \end{bmatrix}$$

The corresponding matrix, with each entry complex-conjugated, is denoted by $\mathbf{X}^*$. Here, $M < N$ is the length of the FIR filters $\mathbf{h}_1$ and $\mathbf{h}_2$ modeling the linear and conjugated channel responses. Furthermore, to allow the modeling of the additional memory effects due to delay errors in the RF cancellation signal, as well as possible time misalignment between $\mathbf{y}_{ADC}$ and $\mathbf{X}$, the system is made non-causal, or delayed, by the removal of $K < M$ upper rows from $\mathbf{X}$ and $\mathbf{X}^*$ [41]. In essence, this means that the first K taps of $\mathbf{h}_1$ and $\mathbf{h}_2$ represent a non-causal or pre-cursor part of the impulse

responses. Based on this, the augmented convolution matrix $\mathbf{X}_{aug}$ can be written as

$$\mathbf{X}_{aug} = \begin{bmatrix} \mathbf{X} & \mathbf{X}^* \end{bmatrix} = \begin{bmatrix} x(M+K-1) & \cdots & x(K) & x^*(M+K-1) & \cdots & x^*(K) \\ x(M+K) & \cdots & x(K+1) & x^*(M+K) & \cdots & x^*(K+1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ x(N-1) & \cdots & x(N-M) & x^*(N-1) & \cdots & x^*(N-M) \end{bmatrix}$$

while the augmented channel $\mathbf{h}_{aug}$ contains $\mathbf{h}_1$ and $\mathbf{h}_2$ stacked as $\mathbf{h}_{aug} = \begin{bmatrix} \mathbf{h}_1^T & \mathbf{h}_2^T \end{bmatrix}^T$.

Using the previous notation, the least-squares estimator for the augmented channel $\mathbf{h}_{aug}$ can then be calculated as

$$\hat{\mathbf{h}}_{aug} = (\mathbf{X}_{aug}^H \mathbf{X}_{aug})^{-1} \mathbf{X}_{aug}^H \mathbf{y}_{ADC}, \tag{46}$$

where $(\cdot)^H$ denotes the Hermitian transpose. In practice, alternative computationally efficient and numerically stable methods, such as singular value decomposition (SVD) based computations, can be used to evaluate the pseudo-inverse in (46). From the augmented channel estimate, due to the stacking property, the individual channel estimates for the direct and image channel can then be obtained directly as

$$\hat{\mathbf{h}}_1 = \begin{bmatrix} \hat{\mathbf{h}}_{aug}(0) & \hat{\mathbf{h}}_{aug}(1) & \cdots & \hat{\mathbf{h}}_{aug}(M-1) \end{bmatrix}^T$$
$$\hat{\mathbf{h}}_2 = \begin{bmatrix} \hat{\mathbf{h}}_{aug}(M) & \hat{\mathbf{h}}_{aug}(M+1) & \cdots & \hat{\mathbf{h}}_{aug}(2M-1) \end{bmatrix}^T$$

Finally, the actual WL cancellation coefficients are directly assigned through $w_1(n) = \hat{h}_1(n)$ and $w_2(n) = \hat{h}_2(n)$, as is evident from (44). With good estimates for the direct and conjugated channels, both the linear SI and conjugate SI will be efficiently attenuated, as will be illustrated in Section IV.

IV. Performance Simulations and Examples

The performance of the proposed WL digital SI cancellation method, including the WL least-squares parameter estimation, is next assessed and illustrated using full-scale waveform simulations of a complete full-duplex transceiver during simultaneous transmission and reception. The waveform simulator is implemented with Matlab, where each component is modeled explicitly and actual OFDM signals are used to ensure that also the signals themselves are realistic. The overall structure of the simulated transceiver is similar to the model presented in Fig. 1. The nonlinearity of the PA is realized by modeling it with a memory polynomial, whose coefficients are derived based on the intercept points. This means that, when the signal is amplified, it will also be distorted in a realistic manner. The effect of IQ imbalance is modeled by introducing some amplitude and phase errors to the I- and Q-branches. These errors are chosen so that they produce the desired value of image attenuation. Also the analog-to-digital conversion is modeled explicitly as a uniform quantization process, which means that the effect of quantization noise is incorporated into the simulations. The effect of thermal noise is realized by summing a normally distributed random signal to the overall signal at the receiver input, after which its power is increased according to the component gains and noise figures. Note that, due to the realistic and detailed modeling of the transceiver chain, no simplifications are made in the simulations regarding the different nonidealities.

TABLE II
Additional parameters for the full-duplex waveform simulator. Baseline parameters are listed in Table I.

Parameter	Value
Constellation	16-QAM
Number of subcarriers	64
Number of data subcarriers	48
Guard interval	25 % of symbol length
Sample length	15.625 ns
Symbol length	4 μs
Oversampling factor	4

Essentially the same parameters are used in the simulations as in the system calculations example, presented in Table I. The additional parameters, describing the utilized OFDM waveform, are presented in Table II. These parameters are in essence similar to WLAN specifications, and both the transmitted and received signals are generated according to them. Similar to the system calculations example, due to the chosen power level for the received signal, SNR at the input of the detector is 15 dB when there are no non-idealities in addition to thermal noise. This serves as the reference value for the measured SINR in assessing how well the total SI can be suppressed in the receiver chain. Namely, if the SI is perfectly cancelled, a SINR of 15 dB is achieved.

In an individual realization, the SI coupling channel between the antennas is assumed to be static, and it is modelled as a line-of-sight component and two weak multipath components, delayed by one and two sample intervals. The average ratio between the power of the main component and the total power of the multipath components is chosen to be 35.8 dB, which is a realistic value for a SI coupling channel when the full-duplex transceiver is located indoors [7].

In the simulations, RF cancellation attenuates only the direct component, as assumed also, e.g., in [1]–[3], whereas the attenuation of the weaker multipath components is done by the proposed digital cancellation algorithm after the ADC. In addition, some delay and amplitude errors are included in the RF cancellation signal to achieve the desired amount of SI attenuation, and to model the cancellation process in a realistic manner. The delay error is implemented as a fractional delay for the cancellation signal, whose value is set to roughly 10 % of sample duration. Together with a small amplitude error, the amount of achieved RF cancellation can then be set to the desired value.

The proposed WL cancellation scheme is analyzed based on two metrics: the SINR at detector input, and the amount of achieved digital SI attenuation. In the simulations, the actual SI channel estimation is done during a specified training period, where there is no actual received signal of interest present. This corresponds to the analysis presented earlier in this paper. The hereby obtained cancellation filter coefficients are then applied in the actual receiver operation mode, with useful received signal present, where the proposed WL canceller output signal, with estimated coefficients, is subtracted from the total received signal. This naturally allows then also the measurement of the SINR at detector input.

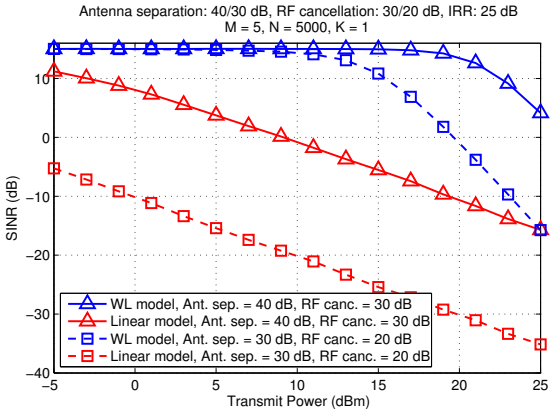


Fig. 5. The SINRs achieved with the proposed WL digital cancellation and with traditional linear digital cancellation. The figures are shown for higher and lower amounts of analog SI attenuation.

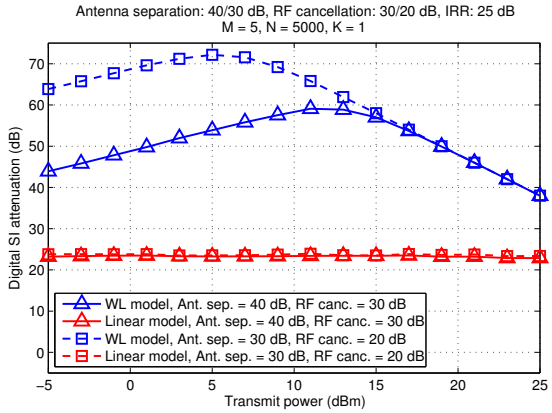


Fig. 6. The amounts of achieved digital cancellation using the proposed WL processing and using traditional linear processing. The figures are shown for higher and lower amounts of analog SI attenuation.

Unlike in Section II-B, the amount of achieved digital SI attenuation is now redefined as the decrease in the power of both the linear SI and conjugate SI. It can thus be written as

$$|a_{WLDC}|^2 = \frac{E[|y_{WLDC,SI}(n)|^2]}{E[|y_{ADC,SI}(n)|^2]} = \frac{E[|(h_1(n) - w_1(n)) \star x(n) + (h_2(n) - w_2(n)) \star x^*(n)|^2]}{E[|h_1(n) \star x(n) + h_2(n) \star x^*(n)|^2]} \quad (47)$$

This definition for digital SI attenuation is used in order to be able to compare the performances of classical linear and widely-linear models for digital cancellation. In the simulations, the amount of digital SI attenuation is then measured according to (47), using the coefficients that were estimated under regular noisy conditions. In these performance measurements, no simplifications are done in order to make sure that the obtained cancellation figures are as accurate as possible.

In the first simulations, transmit power is varied with 2 dB intervals, and the simulation is repeated for 1000 independent realizations for each transmit power. The average value of these runs for the SINR and the amount of digital cancellation is then used in the figures. The number of training samples used for cancellation parameter estimation (N) is fixed to 5000, and the value of K is set to 1, as the fractional delay of the RF cancellation signal may produce additional non-causal memory effects to the resulting signal [41]. The length of the channel estimate (M) is set to 5, as it will allow the modeling of the multipath components and some of the additional memory effects of the SI channel.

In Fig. 5, the SINR with respect to transmit power is shown for both the proposed WL digital canceller, and for the traditional linear digital cancellation method. The latter corresponds to a situation, where the channel estimate is calculated with linear least squares as $\hat{\mathbf{h}}_1 = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{y}_{ADC}$, and $\hat{\mathbf{h}}_2 = \mathbf{0}$. The SINR curves are shown for two cases: in the first case, the amounts of antenna separation and RF cancellation are 30 dB and 40 dB, whereas in the second case they are only 30 dB and 20 dB, respectively.

When investigating the curves corresponding to the higher amount of analog SI attenuation (40 dB of antenna separation and 30 dB of RF cancellation) in Fig. 5, it can be observed that the SINR declines steadily when using only the classical linear model and linear SI cancellation. The effect of the conjugate SI increases with higher transmit powers, as its power is directly related to the power of the SI signal before digital cancellation. On the other hand, when using the proposed WL model and WL processing, it can be observed that the SINR remains essentially at the ideal level of 15 dB with transmit powers below 15 dBm. After this point, it is not possible to achieve the ideal SINR even with the WL digital canceller, as the SINR is decreased by the PA-induced IMD, whose power is increasing rapidly with higher transmit powers. Another factor contributing to the decrease in the SINR is the decreasing resolution of the signal of interest, caused by quantization noise. This is because ideal automatic gain control is assumed in the receiver, always matching the total ADC input waveform to the available ADC voltage range. Nevertheless, a significant improvement in the SINR can be achieved when the proposed WL digital cancellation method is used and also conjugate SI is attenuated in the digital domain.

Since it might be in some instances quite optimistic to assume 40 dB of antenna separation and 30 dB of RF cancellation, in Fig. 5 the SINRs are also shown for a situation where both antenna separation and RF cancellation are 10 dB lower. This will mean that the power of the total SI signal with respect to the signal of interest is significantly higher in the digital domain. Now it can be observed that, when using WL digital cancellation, the residual SI starts to degrade the SINR already with transmit powers above 7 dBm. This is caused by PA-induced IMD, which cannot be modeled by the proposed WL canceller, as discussed earlier. In addition, at this point quantization noise is also slightly decreasing the SINR. However, again, the improvement over traditional linear canceller is significant, as the SINR for the linear model is -5 dB already with the lowest considered transmit power,

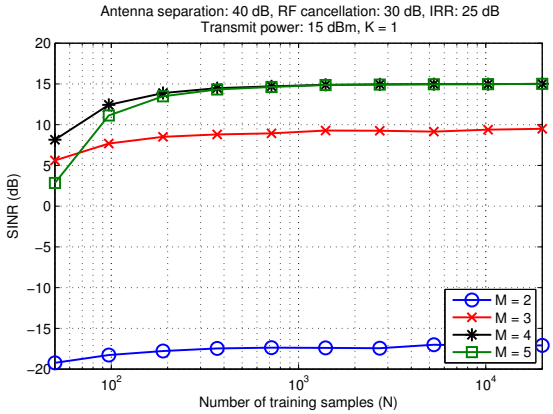


Fig. 7. The SINRs achieved with different lengths for the channel estimate filters, with respect to the number of training samples, using WL digital cancellation.

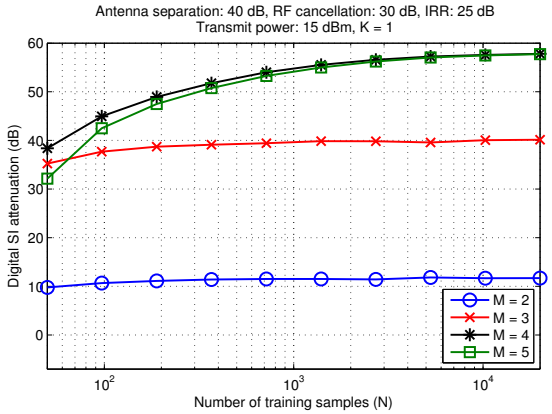


Fig. 8. The amount of digital SI attenuation achieved with different lengths for the channel estimate filters, with respect to the number of training samples, using WL digital cancellation.

when assuming this amount of analog SI attenuation. Thus, the significance of using the proposed WL model for digital cancellation is further emphasized when having less analog SI attenuation. These observations also motivate towards joint cancellation of linear SI, conjugate SI as well as nonlinear SI due to PA. This forms the topic of our future work.

Figure 6 shows the amount of achieved digital SI attenuation for WL processing and traditional linear processing, corresponding to the SINR curves shown in Fig. 5. It can be observed that the amount of achieved digital SI attenuation with the linear model is below 25 dB, regardless of the amount of analog SI attenuation. On the other hand, the achieved SI attenuation with the WL model is as much as 35 dB or 50 dB higher, depending on the amount of antenna separation and RF cancellation. This is explained by the fact that, when using the linear model, the power of the conjugate SI is in essence the noise floor in the digital domain, and it determines the lowest achievable total SI power. Thus, the amount of achieved digital SI attenuation is limited by the power of the conjugate SI. For the WL model, however, this is not the case, and significantly higher values of SI cancellation can be achieved, limited only by the accuracy of the channel estimate.

It can also be observed from Fig. 6 that when the amount of analog SI attenuation is higher, the amount of achieved digital SI attenuation is lower, assuming that the transmit power is below 15 dBm. This is explained by the SNR of the total SI signal in the two scenarios: when the amount of analog SI attenuation is low, the power of the total SI signal with respect to the noise is higher in the digital domain, resulting in a more accurate channel estimate. This obviously means that more digital SI attenuation is achieved when the amount of analog SI attenuation is lower, which has also been observed in actual RF measurements [5].

However, with transmit powers above 15 dBm, the amount of achievable digital SI attenuation is also limited by the IMD, as the nonlinear PA is hindering the parameter estimation accuracy at higher transmit powers. More specifically, the nonlinear distortion due to PA causes a bias to the estimates, as shown in the Appendix, and decreases their accuracy. Thus, with higher transmit powers, the performance of the WL digital cancellation decreases due to increasing bias in the estimation, as well as due to higher residual SI power caused by the IMD. Furthermore, according to (11)–(14), IMD is attenuated in the analog domain by the same amount as the linear SI and conjugate SI signals, and hence its relative power level in the digital domain does not depend on the amount of analog SI attenuation. Thus, with higher transmit powers, when IMD is the limiting factor, same amount of digital SI attenuation is achieved with both sets of values for antenna separation and RF cancellation, as is evidenced by Fig. 6.

Next, the used transmit power is fixed to 15 dBm, and WL model is employed in digital cancellation. The simulation is then run with different lengths of channel estimate filters, and with varying number of training samples, the amount of analog SI attenuation being the same in each case. The number of training samples is varied from 50 to 20000 with 10 logarithmically spaced points in between, and channel estimate filter lengths from 2 to 5 are considered. The value of K is again set to 1 in all cases. The simulation is then repeated for 400 independent runs at each of these points, and the values are averaged.

The SINRs obtained from these simulations are shown in Fig. 7. It can be observed that if the value for M is higher than or equal to 4, there are no significant differences in the achieved SINRs, assuming that the number of training samples is at least 1000. The likely reason for this is that when $M \geq 4$, all the multipath components can be modelled by the channel estimate. This obviously suggests that the channel conditions must be carefully analyzed when choosing a suitable length for the channel estimate filter. Furthermore, the delay error of the RF cancellation signal does not seem to produce any significant causal taps to the total SI channel, as there are no discernible differences in the maximum achievable SINRs between $M = 4$ and $M = 5$.

Figure 7 also indicates that increasing the number of training samples is not likely to bring any benefits after a certain point, as the power of the total SI signal, consisting of both the linear and conjugate SI, is then attenuated well below the noise floor. With the chosen parameters, the value of the SINR saturates with approximately 3000 training samples, regardless of the length of the channel estimate filter. However, it can be observed that the length of the filter affects the minimum number of training samples required to achieve the saturation value of the SINR; with smaller values of M, less samples are required to achieve the highest possible SINR. The reason for this is the higher variance caused by estimating a larger number of coefficients using the same number of training samples.

The amount of achieved digital SI attenuation, corresponding to the SINRs of Fig. 7, is shown in Fig. 8. The form of the curves is similar to the achieved SINRs, i.e., the achieved cancellation is lower with less training samples. Also, the maximum amount of digital SI attenuation is achieved when $M \geq 4$, indicating that this channel estimate length is sufficient to model the actual RF propagation channel. However, this result is heavily dependent on the characteristics of the coupling channel between the antennas, and the accuracy of the delay matching of the RF cancellation signal. Under certain circumstances, longer filters might be needed to model the total SI channel with sufficient accuracy.

In addition, with this transmit power, the amount of digital SI attenuation can be observed to saturate to a value of approximately 58 dB, assuming that a sufficiently high number of training samples is used ($N > 10000$). This saturated value is mainly set by the IMD produced by the PA, which is caused by the fact that the PA distorts the SI waveform in a nonlinear manner, which cannot be modeled by the WL least squares estimation. This, on the other hand, will cause errors in the coefficient estimates, resulting in a lower amount of achieved digital cancellation, as was already discussed earlier. Nevertheless, as was observed from Fig. 7, the highest possible SINR is achieved already with 1000–3000 training samples, depending on the value of M. Thus, based on Fig. 8, it can be concluded that, with transmit powers up to 15 dBm, approximately 55 dB of digital SI attenuation is sufficient to suppress the total SI signal well below the noise floor. Furthermore, our findings indicate that WL processing is required to be able to achieve digital SI attenuation of this magnitude.

V. Conclusion

In this paper, we have proposed a novel method for compensating the image component of the self-interference (SI) signal, caused by IQ imbalances in the transmitter and receiver IQ mixers. To the best of our knowledge, this is the first time this problem has been addressed in the literature at self-interference waveform and digital cancellation levels. With fundamental system calculations, it was first shown that the image component of the SI signal will affect the performance of a full-duplex transceiver by significantly decreasing the maximum achievable SINR. Then, by using the developed widely-linear least-squares parameter estimation and the proposed widely-linear digital cancellation, we showed that this image component can also be attenuated in the digital domain, and thus the decrease in the SINR can be prevented. These claims were affirmed with extensive full waveform simulations, which demonstrate that the proposed method will significantly improve the performance of a typical full-duplex transceiver. Future work will focus on developing joint augmented nonlinear digital cancellation processing that is able to suppress the classical linear SI, nonlinear SI produced by the transmitter power amplifier, and their image components due to IQ imaging.

Appendix
Analysis of Bias

Direct substitution of (45) into (46) yields $\hat{\mathbf{h}}_{aug} = \mathbf{h}_{aug} + (\mathbf{X}_{aug}^H \mathbf{X}_{aug})^{-1} \mathbf{X}_{aug}^H \mathbf{z} = \mathbf{h}_{aug} + \mathbf{e}$. To analyze the estimation error $\mathbf{e}$ analytically, we resort to the simplified frequency-independent model presented in Subsection II-B, that is $y_{ADC}(n) \approx h_1 x(n) + h_2 x^*(n) + h_{IMD} x_{IMD}(n) + u(n)$, where $x_{IMD}(n) = x_{IQ}^{TX}(n)|x_{IQ}^{TX}(n)|^2$ and $x_{IQ}^{TX}(n)$ is as defined in (1), implying that $\mathbf{z} = h_{IMD}\mathbf{x}_{IMD} + \mathbf{u}$. Furthermore, with relatively large sample size N, $\mathbf{X}_{aug}^H \mathbf{X}_{aug}$ can be approximated with ensemble augmented covariance matrix $\mathbf{R}_{x,aug}$ of transmit data $x(n)$ [37], that is

$$\mathbf{R}_{x,aug} = \begin{bmatrix} r & c^* \\ c & r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \approx \frac{1}{N} \mathbf{X}_{aug}^H \mathbf{X}_{aug},$$

where $r = E[|x(n)|^2]$ and $c = E[x(n)^2] = 0$ assuming $x(n)$ is 2nd-order circular. Hence, the estimation error can be written as

$$\mathbf{e} \approx h_{IMD} N^{-1} \mathbf{R}_{x,aug}^{-1} \mathbf{X}_{aug}^H \mathbf{x}_{IMD} + N^{-1} \mathbf{R}_{x,aug}^{-1} \mathbf{X}_{aug}^H \mathbf{u}$$

Due to the dependence of $\mathbf{X}_{aug}$ and $\mathbf{x}_{IMD}$, the average estimation error is nonzero, i.e., the estimator is biased. Assuming that $x(n)$ is also 4th-order circular (i.e., $E[x(n)^4] = E[x(n)^3 x^*(n)] = 0$), this can be shown directly as

$$\begin{aligned} & E[h_{IMD} N^{-1} \mathbf{R}_{x,aug}^{-1} \mathbf{X}_{aug}^H \mathbf{x}_{IMD}] \\ & = h_{IMD} N^{-1} \mathbf{R}_{x,aug}^{-1} E[\mathbf{X}_{aug}^H \mathbf{x}_{IMD}] \\ & \approx h_{IMD} N^{-1} \mathbf{R}_{x,aug}^{-1} \begin{bmatrix} g_{1,TX}|g_{1,TX}|^2 N E[|x(n)|^4] \\ 2g_{2,TX}|g_{1,TX}|^2 N E[|x(n)|^4] \end{bmatrix} \\ & = \frac{h_{IMD}|g_{1,TX}|^2 E[|x(n)|^4]}{r} \begin{bmatrix} g_{1,TX} \\ 2g_{2,TX} \end{bmatrix} \\ & \neq \mathbf{0} \text{ q.e.d.} \end{aligned}$$

Thus, as the above derivation shows, the IMD of the nonlinear transmitter power amplifier causes a bias to the estimator.

References

[1] B. Radunovic, D. Gunawardena, P. Key, A. Proutiere, N. Singh, V. Balan, and G. DeJean, "Rethinking indoor wireless mesh design: Low power, low frequency, full-duplex," in *Proc. Fifth IEEE Workshop on Wireless Mesh Networks*, Jun. 2010, pp. 1–6.

[2] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel full duplex wireless communication," in *Proc. 16th Annual International Conference on Mobile Computing and Networking*, Sep. 2010, pp. 1–12.

[3] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. 17th Annual International Conference on Mobile computing and Networking*, Sep. 2011, pp. 301–312.

[4] A. Sahai, G. Patel, and A. Sabharwal, "Pushing the limits of full-duplex: Design and real-time implementation," Department of Electrical and Computer Engineering, Rice University, Technical Report TREE1104, Jul. 2011.

[5] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. 44th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2010, pp. 1558–1562.

[6] B. Day, A. Margetts, D. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3702–3713, Jul. 2012.

[7] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[8] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. Ramakrishnan, C. Rice, and N. Shankaranarayanan, "Design and characterization of a full-duplex multiantenna system for WiFi networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 3, pp. 1160–1177, Mar. 2014.

[9] M. E. Knox, "Single antenna full duplex communications using a common carrier," in *Proc. IEEE 13th Annual Wireless and Microwave Technology Conference*, Apr. 2012, pp. 1–6.

[10] C. Cox and E. Ackerman, "Demonstration of a single-aperture, full-duplex communication system," in *Proc. Radio and Wireless Symposium*, Jan. 2013, pp. 148–150.

[11] N. Phungamngern, P. Uthansakul, and M. Uthansakul, "Digital and RF interference cancellation for single-channel full-duplex transceiver using a single antenna," in *Proc. 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, May 2013, pp. 1–5.

[12] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[13] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, "MIDU: enabling MIMO full duplex," in *Proc. 18th Annual International Conference on Mobile Computing and Networking*, Aug. 2012, pp. 257–268.

[14] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *SIGCOMM'13*, Aug. 2013.

[15] J. McMichael and K. Kolodziej, "Optimal tuning of analog self-interference cancellers for full-duplex wireless communication," in *Proc. 50th Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2012, pp. 246–251.

[16] J.-H. Lee, "Self-interference cancelation using phase rotation in full-duplex wireless," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4421–4429, Nov. 2013.

[17] L. Anttila, D. Korpi, V. Syrjälä, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*, Nov. 2013, pp. 1193–1198.

[18] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Self-interference cancellation with nonlinear distortion suppression for full-duplex systems," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*, Nov. 2013, pp. 1199–1203.

[19] T. Riihonen, M. Vehkapera, and R. Wichman, "Large-system analysis of rate regions in bidirectional full-duplex MIMO link: Suppression versus cancellation," in *Proc. 47th Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2013, pp. 1–6.

[20] T. Riihonen, P. Mathecken, and R. Wichman, "Effect of oscillator phase noise and processing delay in full-duplex ofdm repeaters," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 1947–1951.

[21] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "Understanding the impact of phase noise on active cancellation in wireless full-duplex," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 29–33.

[22] V. Syrjälä, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," *IEEE Transactions on Wireless Communications*, 2014.

[23] A. Sahai, G. Patel, and A. Sabharwal, "Asynchronous full-duplex wireless," in *Proc. Fourth International Conference on Communication Systems and Networks (COMSNETS)*, Jan. 2012, pp. 1–9.

[24] E. Ahmed, A. Eltawil, and A. Sabharwal, "Rate gain region and design tradeoffs for full-duplex wireless communications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, pp. 3556–3565, Jul. 2013.

[25] T. Riihonen and R. Wichman, "Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 45–49.

[26] S. Li and R. Murch, "Full-duplex wireless communication using transmitter output based echo cancellation," in *Proc. Global Telecommunications Conference (GLOBECOM)*, Dec. 2011, pp. 1–5.

[27] D. Bliss, T. Hancock, and P. Schniter, "Hardware phenomenological effects on cochannel full-duplex MIMO relay performance," in *Proc. 46th Asilomar Conference on Signals, Systems and Computers*, Nov. 2012, pp. 34–39.

[28] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2498–2511, May 2013.

[29] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Transactions on Wireless Communications*, 2014.

[30] L. Anttila, "Digital front-end signal processing with widely-linear signal models in radio devices," Ph.D. dissertation, Tampere University of Technology, 2011.

[31] "LTE; evolved universal terrestrial radio access (E-UTRA); user equipment (UE) radio transmission and reception (3GPP TS 36.101 version 11.2.0 release 11)," ETSI, Sophia Antipolis Cedex, France.

[32] Y. Hua, P. Liang, Y. Ma, A. Cirik, and Q. Gao, "A method for broadband full-duplex MIMO radio," *Signal Processing Letters*, vol. 19, no. 12, pp. 793–796, Dec. 2012.

[33] "MAX2829 single-/dual-band 802.11a/b/g world-band transceiver IC," Maxim Integrated, San Jose, California, USA.

[34] Q. Gu, *RF System Design of Transceivers for Wireless Communications*. Springer-Verlag New York, Inc., 2006.

[35] A. Pärssinen, J. Jussila, J. Ryynänen, L. Sumanen, and K. A. I. Halonen, "A 2-GHz wide-band direct conversion receiver for WCDMA applications," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 12, pp. 1893–1903, Dec. 1999.

[36] H. Yoshida, T. Kato, T. Toyoda, I. Seto, R. Fujimoto, T. Kimura, O. Watanabe, T. Arai, T. Itakura, and H. Tsurumi, "Fully differential direct conversion receiver for W-CDMA using an active harmonic mixer," in *Proc. Radio Frequency Integrated Circuits Symposium*, Jun. 2003, pp. 395–398.

[37] P. Schreier and L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge University Press, 2010.

[38] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.

[39] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Transactions on Communications*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[40] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[41] O. Hammi, F. Ghannouchi, and B. Vassilakis, "On the sensitivity of RF transmitters' memory polynomial model identification to delay alignment resolution," *Microwave and Wireless Components Letters*, vol. 18, no. 4, pp. 263–265, Apr. 2008.

Dani Korpi was born in Ilmajoki, Finland, on November 16, 1989. He received the B.Sc. and M.Sc. degrees (both with honours) in communications engineering from Tampere University of Technology (TUT), Finland, in 2012 and 2014, respectively.

He is currently a researcher at the Department of Electronics and Communications Engineering at TUT, pursuing the D.Sc. (Tech.) degree in communications engineering. His main research interest is the study and development of single-channel full-duplex radios, with a focus on analysing the RF impairments.

Lauri Anttila (S'06, M'11) received the M.Sc. (Tech.) degree in 2004 and the D.Sc. (Tech.) degree (with honours) in 2011 in communications engineering from Tampere University of Technology (TUT), Tampere, Finland.

Currently, he is a Research Fellow at the Department of Electronics and Communications Engineering at TUT. His general research interests include statistical and adaptive signal processing for communications, digital front-end signal processing in exible radio transceivers, radio architectures, and full-duplex radio systems.

Ville Syrjälä (S'09, M'12) was born in Lapua, Finland, in 1982. He received the M.Sc. (Tech.) degree in 2007 and D.Sc. (Tech.) degree in 2012 in communications engineering (CS/EE) from Tampere University of Technology (TUT), Finland.

He was working as a research fellow with the Department of Electronics and Communications Engineering at TUT, Finland, until 2013. Currently, he is working as a research fellow of the Japan Society for the Promotion of Science (JSPS) at Kyoto University, Japan. His general research interests are in full-duplex radio technology, communications signal processing, transceiver impairments, signal processing algorithms for flexible radios, transceiver architectures, direct sampling radios, and multicarrier modulation techniques.

Mikko Valkama (S'00, M'02) was born in Pirkkala, Finland, on November 27, 1975. He received the M.Sc. and Ph.D. degrees (both with honours) in electrical engineering (EE) from Tampere University of Technology (TUT), Finland, in 2000 and 2001, respectively. In 2002 he received the Best Ph.D. Thesis award by the Finnish Academy of Science and Letters for his dissertation entitled "Advanced I/Q signal processing for wideband receivers: Models and algorithms".

In 2003, he was working as a visiting researcher with the Communications Systems and Signal Processing Institute at SDSU, San Diego, CA. Currently, he is a Full Professor and Department Vice Head at the Department of Electronics and Communications Engineering at TUT, Finland. He has been involved in organizing conferences, like the IEEE SPAWC07 (Publications Chair) held in Helsinki, Finland. His general research interests include communications signal processing, estimation and detection techniques, signal processing algorithms for software defined flexible radios, full-duplex radio technology, cognitive radio, digital transmission techniques such as different variants of multicarrier modulation methods and OFDM, radio localization methods, and radio resource management for ad-hoc and mobile networks.

Publication 3

M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodríguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full-duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, May 2015. DOI: 10.1109/MCOM.2015.7105647

Recent Advances in Antenna Design and Interference Cancellation Algorithms for In-band Full-Duplex Relays

Mikko Heino, Dani Korpi, Timo Huusari, Emilio Antonio-Rodríguez, Sathya Venkatasubramanian, Taneli Riihonen, Lauri Anttila, Clemens Icheln, Katsuyuki Haneda, Risto Wichman, and Mikko Valkama

***Abstract*—In-band full-duplex relays transmit and receive simultaneously at the same center frequency, hence offering enhanced spectral efficiency for relay deployments. In order to deploy such full-duplex relays, it is necessary to efficiently mitigate the inherent self-interference stemming from the strong transmit signal coupling to the sensitive receive chain. In this article, we present novel state-of-the-art antenna solutions as well as digital self-interference cancellation algorithms for compact multiple-input multiple-output (MIMO) full-duplex relays, specifically targeted for reduced-cost deployments in local area networks. The presented antenna design builds on resonant wavetraps and is shown to provide passive isolations in the order of 60–70 dB. We also discuss and present advanced digital cancellation solutions, beyond classical linear processing, specifically tailored against nonlinear distortion of the power amplifier when operating close to saturation. Measured results from a complete demonstrator system, integrating antennas, RF cancellation and nonlinear digital cancellation, are also presented, evidencing close to 100 dB of overall self-interference suppression. The reported results indicate that building and deploying compact full-duplex MIMO relays is already technologically feasible.**

I. Introduction

THE industry envisions that upcoming 5G radio communication systems will provide even thousand-fold throughput compared to the current 4G systems while in-band full-duplex technology [1] may only double the spectral efficiency at best. This may sound modest given the ambitious targets of 5G systems, but somebody with a can-do attitude would think instead that "great, we are already halfway through." Nevertheless, attaining this target requires in any case a combination of several different physical and network layer techniques and without in-band full-duplex technology 5G networks would not reach their full potential.

An in-band full-duplex capable transceiver is able to transmit and receive simultaneously over the same center frequency [1]. With this definition, we can identify three basic communication scenarios of two or three nodes which cover any generic wireless network (such as the one shown at the bottom of Fig. 1) when they are freely combined together. In particular, the possible use cases benefiting from full-duplex operation include bidirectional device-to-device transmission between two full-duplex user terminals, full-duplex access point serving uplink and downlink (possibly half-duplex) users simultaneously in the same radio resource and two-hop full-duplex wireless relaying (as shown in the middle of Fig. 1). Actually, the latter two use cases are topologically the same with only the conceptual difference that relays always forward their received information while the uplink user is not likely communicating with the downlink user but somebody else in a totally different cell. In addition to increasing spectral efficiency, full-duplex capable nodes may decrease latency within multi-hop links and avoid the guard period that is necessary when switching the direction of the transmission from downlink to uplink in time-division duplex (TDD) systems. On network level, full-duplex nodes may simultaneously process user plane and control plane signals decreasing the latency and boosting up the operation of the overall system.

In practice, the increase in throughput due to full-duplex operation is limited by the presence of unavoidable self-interference when the transmitted signal couples back to the receiver in the in-band full-duplex transceiver. Even when the transmitted signal is known in digital baseband, it cannot be eliminated completely in the receiver, because of RF impairments [2] and a large power difference between the transmitted and received signals. On the other hand, RF cancellation is limited due to the simplified processing capabilities of electronics when compared to digital signal processing. Nevertheless, 70–100 dB overall attenuation levels of self-interference have been reported in the literature, e.g. in [3] and [4], and holistic cancellation techniques are currently subject to intensive research by many groups around the world, including our own consortium. Such attenuation levels are already so high that in local area networks, where transmit powers are limited and nodes may be close to each other, co-channel interference may already start to dominate over self-interference, which is our objective as well.

However, even if self-interference is suppressed below receiver noise or ambient co-channel interference, a full-duplex transceiver may outperform its half-duplex counterpart only when there is simultaneous, balanced traffic in both uplink and downlink. Obviously, two-way frequency reuse is useless if data mostly flows in a single direction only. In fact, recent

M. Heino, E. Antonio-Rodríguez, S. Venkatasubramanian, T. Riihonen, C. Icheln, K. Haneda, and R. Wichman are with Aalto University School of Electrical Engineering, Finland.

D. Korpi, T. Huusari, L. Anttila, and M. Valkama are with Tampere University of Technology, Finland, e-mail: mikko.e.valkama@tut.fi

The research work leading to these results was funded by the Academy of Finland (under the project #259915 "In-band Full-Duplex MIMO Transmission: A Breakthrough to High-Speed Low-Latency Mobile Networks"), the Finnish Funding Agency for Technology and Innovation (Tekes, under the project "Full-Duplex Cognitive Radio"), the Linz Center of Mechatronics (LCM) in the framework of the Austrian COMET-K2 programme, and Emil Aaltonen Foundation. The research was also supported by the Internet of Things program of DIGILE (Finnish Strategic Centre for Science, Technology and Innovation in the field of ICT), funded by Tekes.

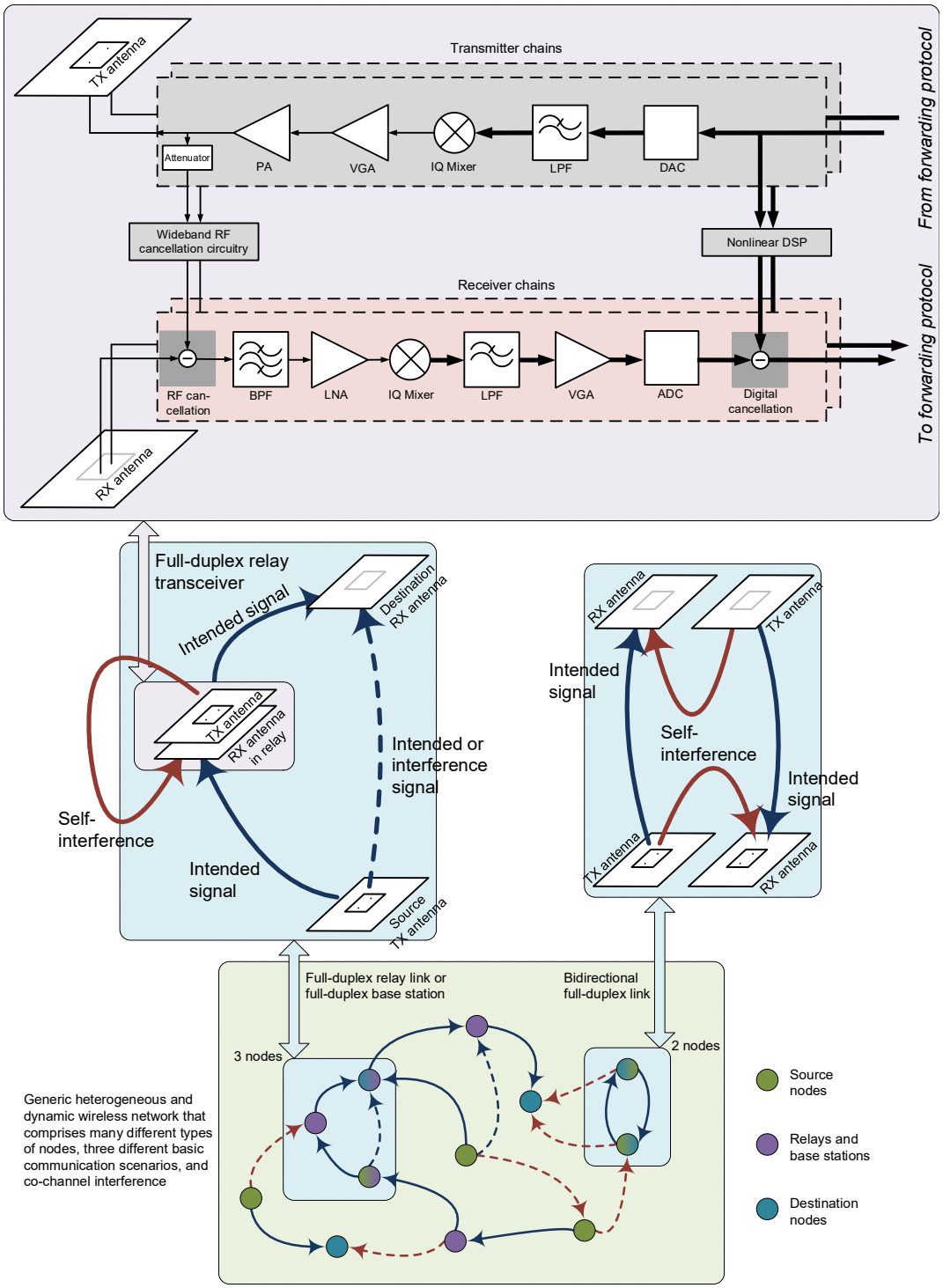


Fig. 1: A multi-antenna radio transceiver and basic two- and three-node full-duplex scenarios occurring in wireless networks.

statistics from cellular network vendors and operators report[1,2] commonly up to 8:2 or 9:1 traffic ratios between downlink and uplink, challenging the applicability of full-duplex access points. Ideal full-duplex operation would yield at best only 25% or 11% throughput improvement with the above traffic ratios, respectively, in comparison to the wanted 100% gain. In contrast, the traffic in full-duplex relaying is inherently symmetric, because relays are supposed to always forward the same amount of information as they receive, making the wireless relays the most prominent use case for in-band full-duplex operation to start with.

In this article, we explore various sophisticated techniques for self-interference mitigation and cancellation within multi-antenna in-band full-duplex relays of compact size suitable for local area networks. By combining antenna design with novel RF and digital cancellation, we show that the self-interference can be suppressed below noise level even when using regular low-cost components. We report measured results from a complete demonstrator system, integrating the advanced antenna designs and analog and digital cancellation, evidencing close to 100 dB of overall isolation between transmitter and receiver chains. Thus, in-band full-duplex relays have the potential to significantly improve the performance in local area networks.

II. Implementation Level Research Challenges in In-band Full-duplex Transceivers

As already discussed, one of the most crucial issues in wireless single channel full-duplex communications is the own transmit signal, or self-interference (SI), that is coupled back to the receiver and acts as a strong source of interference. The SI signal can be as much as 60–100 dB more powerful than the weak received signal of interest, and thereby it must be attenuated significantly to allow the detection of the actual received signal and enable in-band full-duplex communications in the first place. In theory, this is straight-forward as the own transmit signal is known within the device, and thus the SI can in principle be perfectly cancelled merely by subtracting the original transmit signal, properly filtered, from the received signal. In practice, however, achieving a sufficient amount of SI cancellation calls for more advanced and elaborate measures since all the sources of distortion in the transmitter and receiver must be accounted for when subtracting the cancellation signal from the received signal.

A. Passive Isolation and Active Cancellation

To minimize the amount of the more complicated active cancellation techniques, requiring knowledge about the different sources of distortion, the passive isolation between the transmitter and receiver should be maximized. This means that the power of the SI leaking to the receiver is smaller in the first place, and thereby not as much active cancellation is needed to attenuate it to a tolerable level. When separate transmit and receive antennas are used, the electromagnetic isolation between the antennas can be improved in a straightforward manner by increasing the spacing between the antennas or by using different polarizations, resulting in lower SI power. The main limitation in increasing the spacing is that the transceivers are usually space limited, and hence, we do not have the luxury of increasing the distance between the antennas. Moreover, when MIMO is used to improve the link capacity, both polarizations may be used to provide polarization diversity. Hence, advanced methods are required to improve the electromagnetic isolation between the antennas. These methods include, for instance, the use of band-gap structures as high-impedance surfaces [5] to prevent surface waves between the antennas, inductive loops [6] to produce counter-flowing magnetic fields to reduce the coupling between the antennas, and resonant structures like wavetraps [7] and slots on a ground plane [8] to alter ground plane currents to reduce the coupling. Because these techniques are completely passive in nature, they require no tracking of the SI signal and its possible distortion, while still providing a potentially significant increase in the amount of SI attenuation. In addition to the aforementioned methods to improve the antenna isolation, the port-to-port isolation between the antenna feeds can be improved by using techniques like neutralization [9], which improves the isolation using a cancellation path between two antennas, or connecting lumped elements between antenna feeds [10] to cancel the mutual admittance between the antennas, thereby improving the port-to-port isolation. In this article, one possible method of using resonant wavetraps is discussed and demonstrated to improve the electromagnetic isolation in compact devices, especially relays.

It is, however, typically not possible to mitigate the SI signal perfectly with the passive antenna-based techniques, and thus also active SI cancellation is required. A common solution is to do the active attenuation of the SI signal in two stages: first at the input of the receiver chain, operating at RF frequencies, and then after the analog-to-digital conversion. These SI cancellation stages are usually referred to as RF cancellation and digital cancellation, respectively. RF cancellation is required in order to prevent the complete saturation of the receiver components and the analog-to-digital converter (ADC). Finally, digital cancellation is performed to attenuate the remaining SI signal below the noise floor. Both of these active cancellation methods rely on processing the known transmit signal to produce the cancellation signal.

A block diagram of an in-band MIMO full-duplex relay, employing the aforementioned active SI cancellation stages, is illustrated in the upper part of Fig. 1. The transmitter and receiver chains have been chosen to follow direct-conversion architecture, which is a typical selection for modern wireless transceivers. This block diagram depicts the transceiver model that will be utilized in the following chapters, thereby acting as a basis for the forthcoming analysis. In this paper, the emphasis is on a full-duplex relay, meaning that the transmitters and receivers have separate antennas, which can also be seen in Fig. 1. It should be noted, however, that it is in principle also possible to share an antenna between a transmitter and a receiver, even in an in-band full-duplex transceiver [11].

[1]Ericsson, On the Pulse of the Networked Society, Mobility Report, Nov. 2012, available at: http://www.ericsson.com/res/docs/2012/ericsson-mobility-report-november-2012.pdf

[2]Nokia Solutions and Networks, TD-LTE Frame Configuration Primer, white paper, available at: http://networks.nokia.com/system/files/document/nsn_td_lte_frame_configuration_wp.pdf

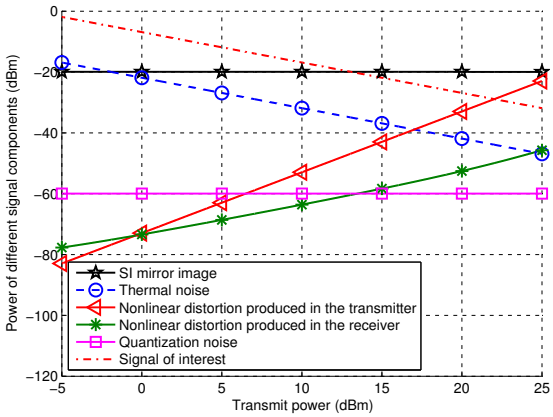


Fig. 2: Example power levels of the different signal component at the input of the receiver detector. In this example, it is assumed that the linear SI component can be perfectly suppressed.

B. RF Imperfections

Typically, the performance of the active SI cancellation mechanisms based on linear processing is limited by the non-idealities occurring within the analog/RF circuit of the full-duplex transceiver [2]. This applies especially to the SI cancellation occurring in the digital domain, where only the original transmit waveform is available. For this reason, the most prominent types of such analog/RF circuit non-idealities must be considered in this context, as they have to be regenerated in the cancellation processing to produce a sufficiently accurate cancellation signal. The most significant impairments include the nonlinear distortion of the SI signal, caused by the different amplifiers in the transmitter and receiver, I/Q imbalance of the transmitter and receiver, phase noise of the local oscillator, and ADC quantization noise. Some of these RF impairments can be modeled on the waveform level, with appropriate behavioral models, within digital cancellation, and thereby they do not pose an insurmountable obstacle for in-band full-duplex communications. Good examples of recent work in this field are [2], [3], [6], [12]–[15]. In particular, the nonlinear distortion of the SI signal can be modeled in the digital domain using a memory polynomial or parallel Hammerstein type of model with carefully chosen coefficients [12], [14]. With such a model, the clean transmit data can be processed to create an accurate SI replica for efficient cancellation. Correspondingly, the effect of I/Q imbalance can be modeled digitally as an additional distortion component, which consists of the complex conjugate of the original baseband transmit signal with certain memory coefficients [13]. Nevertheless, there are also certain sources of distortion that cannot be dealt with afterwards, such as thermal noise and quantization errors produced by the ADC. These must be accounted for already in the design process of the transceiver [2].

To illustrate the need for modeling the most dominant impairments at the waveform level, Fig. 2 shows the absolute power levels of the different signal components at the receiver detector input, with respect to the transmit power. The power levels have been calculated for a SISO in-band full-duplex transceiver using simplified system calculations, similar to [2], and using realistic RF component and ADC specifications. In this type of an analysis, only the power levels of the different signal components are taken into account, which simplifies the calculations considerably while still providing useful insight into the system behavior. In this example, it is assumed that only the linear component of the SI signal can be perfectly attenuated by the different SI cancellation stages. To get the final numerical values for the power levels, typical component parameters from earlier literature and technical specifications have then been used, the most important of which are shown in Table I for readers convenience. Based on Fig. 2, it is evident that at least the transmitter power amplifier induced nonlinearities, alongside with the SI mirror image produced by the I/Q imbalances, must be attenuated by the digital cancellation algorithm to enable reliable in-band full-duplex operation and maintain an acceptable signal-to-interference-plus-noise ratio (SINR). Thus, simple linear modeling and processing in the digital domain does not suffice, and more advanced algorithms are required.

TABLE I: Example component parameters and other essential values used in evaluating the power levels.

Parameter	Value
SNR	15 dB
Bandwidth	12.5 MHz
RX noise figure	4 dB
Antenna isolation	40 dB
RF cancellation	30 dB
TX power amplifier IIP3	13 dBm
I/Q image rejection ratio	30 dB
RX low-noise amplifier IIP3	−9 dBm
Number of bits at the ADC	12

In addition to the analog circuit-level impairments, one additional aspect to consider is the time varying nature of the SI propagation channel. This is often overlooked when discussing SI cancellation in full-duplex devices but in practice it is obviously a key aspect in implementing an in-band full-duplex transceiver. The time varying nature of the SI channel becomes especially crucial in relaying applications where the transmit and receive antennas can be physically separated to achieve higher natural isolation between them. In such scenario, the SI channel can vary significantly in time. However, even in a single-antenna full-duplex device, the reflections from nearby scatterers can change significantly from one moment to the next, thereby warranting the use of adaptive SI cancellation solutions.

When considering a complete and functional in-band full-duplex relay, it is more than likely to involve all of the previously discussed methods to obtain a sufficient amount of SI attenuation. First, the passive isolation between the transmitter and receiver is maximized, after which the SI signal is attenuated actively. In the digital domain, the possible circuit-level impairments are then regenerated with advanced algorithms, and cancelled efficiently. Furthermore, these algorithms also

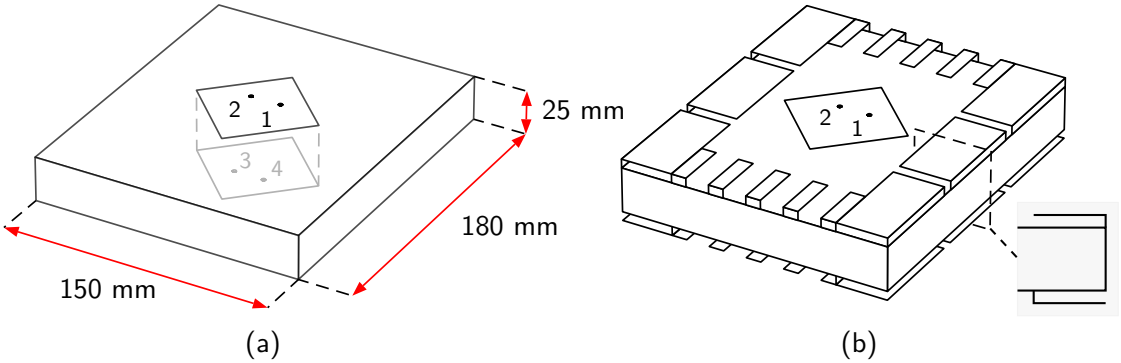


Fig. 3: A compact in-band full-duplex relay antenna (a) without wavetraps (b) with wavetraps. In (b), wavetraps are facing opposite directions on either side of the relay.

have the necessary adaptivity to constantly track the changes in the SI channel parameters. To demonstrate that such a device is quickly becoming a realistic prospect, this article provides some of the latest research results, alongside with measurement data, of these individual key factors, which are necessary to implement a functional in-band full-duplex relay.

III. Advanced Antenna Designs for Enhanced Transmitter-Receiver Isolation

In this section, the use of resonant wavetraps is discussed as a possible method for improving the electromagnetic isolation in a compact device, e.g., a back-to-back relay with transmit and receive antennas on either side of a box as shown in Fig. 3. In Fig. 3a, a closed box of dimensions 180×150 mm with dual polarized patch antennas (50×50 mm) on the top and bottom is shown. The use of a closed box represents a practical scenario with the electronics contained inside the relay. The patch antenna on the bottom side is co-located with the top side patch antenna, but with the antenna feed locations on the opposite sides of the patch (see Fig. 3a). The patch antennas are dual polarized in order to support MIMO. The relay can be used, for example, as an outdoor-indoor relay as discussed in [6]. In order to enable in-band full-duplex operation, the electromagnetic isolation between the antennas on opposite sides of the box has to be improved. Fig. 3b shows the schematic of the relay antenna with wavetraps connected to the ground plane.

Wavetraps are resonant structures that can be used to control the surface currents on the ground planes. In this case, a planar quarter-wavelength patch short-circuited at the other end to the ground plane is used as a wavetrap. According to transmission line theory, the input impedance of a short-circuited quarter-wave transmission line is very high. This results in small currents at the open end of the patch with large currents concentrated within the patch. This reduces the surface currents and the electromagnetic fields on the other side of the relay box, thereby improving the isolation between the upper antenna and the lower antenna of the box.

In this work, the patch antennas of the relay are resonant at 2.56 GHz with a -10 dB impedance bandwidth of 139 MHz and a -6 dB bandwidth of 238 MHz. They can be used, e.g., for WiMAX or LTE. In the relay structure, ports 1 and 2 (TX) are on the top and ports 3 and 4 (RX) on the bottom side, as shown in Fig. 3a. The wavetraps are then implemented as quarter-wave patches as shown in Fig. 3b with one end of the patch short-circuited to the ground plane. The dimensions and number of the wavetraps were optimized with EM-field simulations to obtain the best possible isolation over a sufficient frequency band between the TX and RX antenna ports.

The resonant wavetraps on the longer edge have larger width and are closely spaced whereas along the shorter edge, the width is smaller and the wavetraps are more sparsely placed. This way the best minimum isolation defined as the worst isolation among the four different combinations of the RX and TX ports, i.e., S_{31}, S_{41}, S_{32}, and S_{42}, is obtained. The wavetraps on both sides of the relay are similar but are oriented in opposite directions, i.e., the open ends of the wavetraps are located close to the antenna on the TX side and away from the antenna on the RX side as seen in Fig. 3b. The best isolation between the antenna ports was obtained with this particular asymmetry.

Figures 4a and 4b illustrate the simulated average electric field strength at the center cross section parallel to the shorter edge of the box at 2.56 GHz with and without wavetraps, respectively, when port 1 of the antenna is excited. From Fig. 4a, it can be observed that the transmitted electromagnetic wave diffracts strongly around the corners of the relay enclosure and couples to the receiving side antenna. With the wavetraps, as shown in Fig. 4b, the electric field strength is substantially lower on the receiving side of the relay than without the wavetraps, resulting in clearly improved isolation.

The designed relay antenna was manufactured, and then measured in an anechoic chamber. The isolation was evaluated by measuring the S-parameters of the relay antenna as a 4-port network. The measured S-parameters are presented in Fig. 4c and 4d where S_{11} denotes the antenna input matching and S_{21} and S_{43} the coupling between the antenna ports which are on the same side of the box. The remaining S-parameters show the isolation between the antenna ports on opposite sides of the box. In the relay without the wavetraps, the minimum isolation

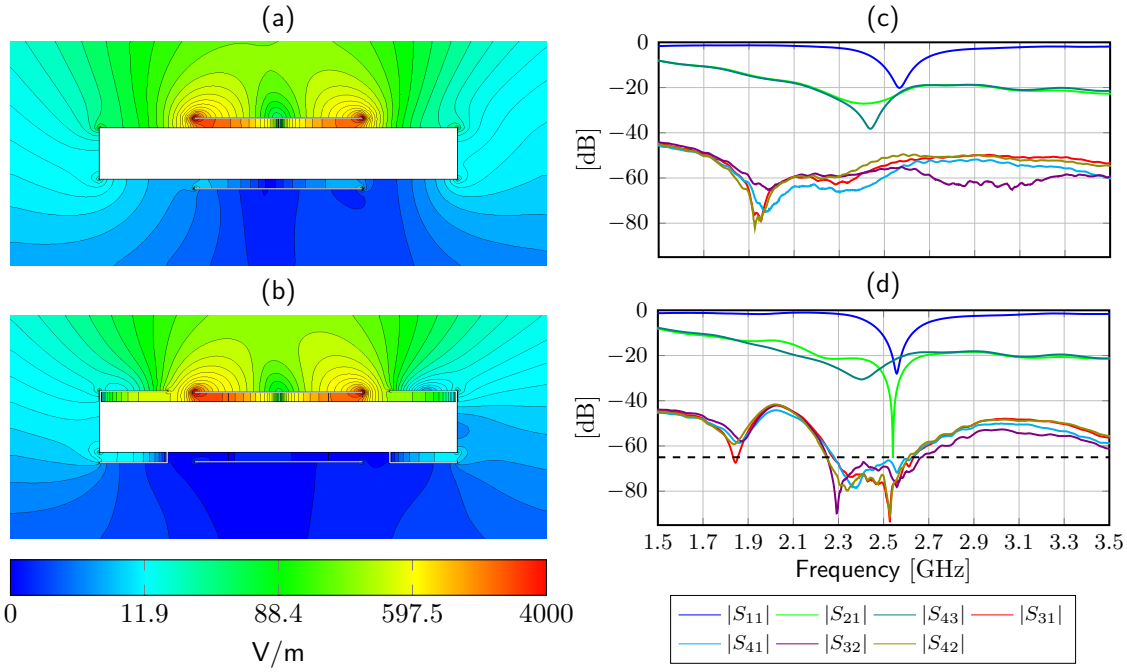


Fig. 4: (a) Simulated electric field strength without wavetraps, (b) with wavetraps, (c) measured S-parameters of the relay without wavetraps and (d) with wavetraps.

between RX and TX ports is 50 dB at the resonance frequency of the patch antenna as shown in Fig. 4c. After adding the wavetraps, the minimum isolation improved by 21 dB to 71 dB at the resonance frequency. In addition, the wavetraps offer also a wideband isolation improvement around the designed frequency, even though they are frequency sensitive. The isolation across a 167 MHz band around the center frequency is better than 65 dB, which is an improvement of 15 dB over the 50 dB minimum isolation without wavetraps as indicated by the dashed horizontal line in Fig. 4d. Within the given isolation bandwidth, the corresponding antenna port matching is better than -6 dB to ensure enough power radiated by the antennas. The gain of the relay antenna was also measured in the anechoic chamber to determine the impact of the wavetraps on the radiation pattern. The impact of adding wavetraps on the ground plane was minor with a maximum gain of 10.2 ± 0.1 dBi for all ports of the relay without wavetraps and 9.6 dBi for ports 1 and 2 and 10.5 ± 0.1 dBi for ports 3 and 4 for the relay with wavetraps. In conclusion, the wavetraps efficiently concentrate the ground plane currents and reduce the coupling on the other side of the relay box, thereby improving the antenna isolation.

IV. Advanced Self-Interference Cancellation Algorithms

Even though antenna isolation and RF cancellation provide a significant level of SI mitigation before digital processing is applied (~60–80 dB), the residual SI after analog-to-digital conversion may still be strong enough to desensitize the receiver [2], [14]. Therefore, the use of additional cancellation techniques in the digital domain is essential in order to reduce the remaining interference below noise levels and guarantee the optimal performance of the system.

A. Basic Principles

As previously explained, cancellation techniques in the digital domain work in a similar way to their analog counterpart, i.e., a digital replica of the SI signal is generated and then subtracted from the received signal. The fidelity of the replica signal will determine the level of residual SI after cancellation and, therefore, will upper bound the system performance [16]. In order to achieve a nearly interference-free system, the selection of a proper cancellation method is of major importance. When both transmit and receive sides of the full-duplex device behave primarily like a linear filter, cancellation reduces to designing a filter that identifies the SI channel, since the transmitted signal is known at every time instant. However, a linear cancellation architecture cannot mitigate nonlinear behavior and noise sources such as nonlinear distortion of the power amplifier, I/Q imbalance during modulation/demodulation, phase noise, or quantization noise at the receiver [2], [3], [13], [14]. Some of these impairments can be mitigated by extending the linear architecture to a nonlinear architecture, as will be described later in this section.

Typically, digital cancellation is performed either in the time domain or in the frequency domain. Cancellation in the frequency domain processes each subcarrier signal individually, which may result in a computationally demanding scheme if the number of subcarriers is large. On the other hand,

cancellation in the time domain processes the signal samples independently of the number of subcarriers, but, due to the different interference paths between antennas, requires gauging the delay spread of the SI channel [17]. As shown in Fig. 1, cancellation takes place after baseband demodulation and digital conversion of the received signal, usually being the first operation within the digital pipeline. As a result, the employed signal is sampled above the Nyquist rate, which demands the use of special techniques to deal with arbitrary signal spectra [17]. Here we focus entirely on time-domain cancellation techniques.

B. Adaptive vs. Block-based Processing

According to the operation mode of the cancellation scheme, we can distinguish between online and offline cancellation. Online cancellation uses an adaptive approach for obtaining the optimal cancellation filter, in which every new sample is employed to iteratively update the cancellation filter. Changes in the environment or in the internal state of the device, such as new coupling paths or reflections, as well as variations in the temperature or the transmitted power, can severely modify the SI channel and, consequently, an updated estimation of the SI channel is essential to sustain optimal interference levels. An online cancellation scheme can track temporal variations of the device environment while maintaining a constant computational load, though the number of samples required to reach an optimal solution might be high [16]. On the other hand, offline cancellation obtains the optimal cancellation filter through a batch operation, after receiving several samples. This approach requires fewer samples to reach a solution, but the tracking capabilities are compromised and the computational load per sample is significantly higher than in online cancellation. A periodic estimation of the SI channel is necessary to maintain good performance levels when considering time-varying scenarios [12]. In general, the number of samples available for cancellation indicates which option to implement. An offline algorithm is preferable when few samples are available, whereas an online algorithm is preferable when a large number of samples is available.

Regardless of the operation mode, the optimal cancellation filter is obtained as the solution of a minimization problem. Assuming that the transmitted signal of the full-duplex device is uncorrelated with the incoming source signal, the optimal cancellation filter minimizes the signal power after cancellation or, equivalently, identifies the SI channel [16]. Consequently, the optimal filter can be obtained using common adaptive techniques like gradient descent or recursive least-squares in the online cancellation case, or using, e.g., conventional least-squares in the offline cancellation case. For detailed algorithm descriptions, refer to [14], [16].

C. Linear vs. Nonlinear Cancellation

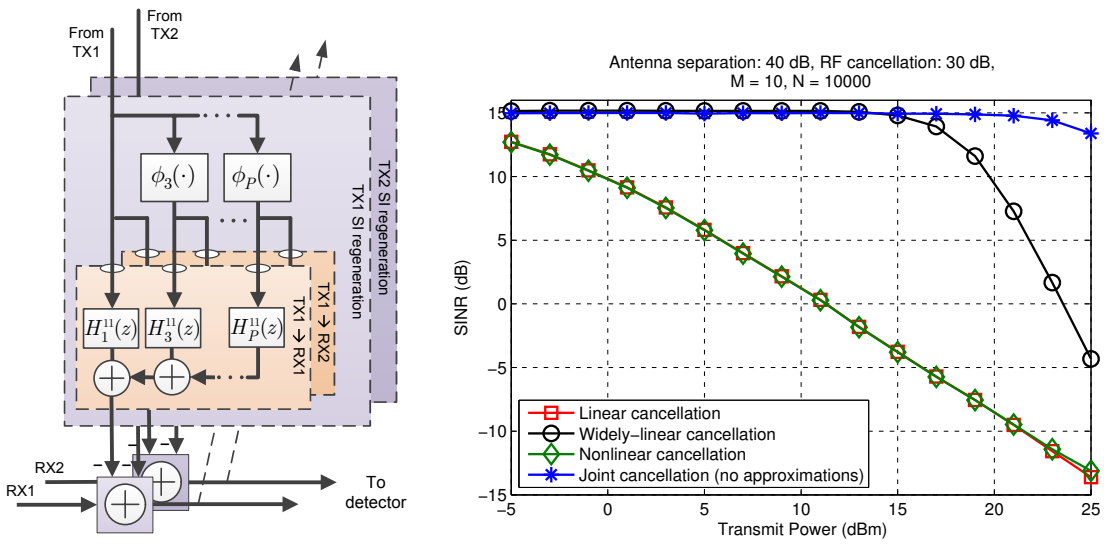
In general, the underlying architecture of the digital cancellation filter can be classified into two types: linear cancellation and nonlinear cancellation. Whereas linear cancellation assumes that the self-interference channel can be modeled as a linear filtering operation, nonlinear cancellation takes into account the presence of nonlinear components in the transmission/reception chains. As it was seen in the system calculations example in Fig. 2, impairments such as power amplifier (PA) nonlinearity and I/Q imbalance may be significant.

Nonlinear effects like amplifier distortion or mixer nonlinearities can be accurately modeled using polynomial-based systems [3], [12], [14], whereas I/Q imbalance can be modeled using widely-linear filters [13], both of which have been extensively studied in the recent literature. In brief, polynomial-based systems model nonlinearities by processing higher-order terms of the input signal, i.e., $\{x_n, x_n|x_n|^2, x_n|x_n|^4, \ldots\}$, when interpreted for complex-valued signals and odd-order nonlinearities [12], [14], whereas widely-linear systems model I/Q imbalance by separately filtering the input signal and its complex conjugate, i.e., $\{x_n, x_n^*\}$ [13].

As a consequence of performing the cancellation in the time domain, the self-interference propagation channel from each transmit antenna to each receive antenna is modeled using tapped delay lines, so the overall nonlinear model for a full-duplex device will consist of a linear combination of polynomial basis functions [14]. In general, in a full-duplex MIMO device with N transmit antennas and M receive antennas, $N \times M$ self-interference channels need to be identified. Note that the number of parameters grows linearly with both N and M and the number of delay taps.

Figure 5a shows a possible nonlinear canceller architecture for the full-duplex device in Fig. 1, assuming two transmit antennas and two receive antennas, i.e., $N = M = 2$, and a nonlinear architecture based on polynomial models. Function $\phi_p(x_n) = x_n|x_n|^{p-1}$ is the polynomial basis function of order p and $H_p(z)$ is the linear filter associated with $\phi_p(x_n)$. Such nonlinear architecture, where a static nonlinearity is followed by a linear filter, is called a parallel Hammerstein model. Hammerstein models are typically used for modeling PA nonlinearities, and have been found to be excellent behavioral models in terms of the accuracy-complexity trade-off [18]. In general, the number of parameters of a Hammerstein model grows linearly with order p, while in the MIMO case, the increase is relative to $N \times M$ [14]. In [12] the parallel Hammerstein model is successfully used to model the overall nonlinear SI channel comprising a nonlinear PA and a multipath SI channel. While [12] treated the single-antenna case, in [14] this model is extended to the MIMO case, where also the I/Q imbalance at the transmit side is taken into account through the use of additional complex-valued basis functions.

To illustrate the potential of the digital cancellation solutions, simulations with the techniques employed in [14] are presented. Figure 5b shows the performance results of different cancellation schemes for a full-duplex device where both transmit and receive signals are 12.5 MHz OFDM signals, and the signal-to-noise ratio at reception, per receiver, is 15 dB. The SI propagation channel between the antennas is a multipath channel, and power amplifier nonlinear distortion and I/Q imbalance during modulation are also included in the system model. Furthermore, an antenna isolation of 40 dB and analog RF cancellation of 30 dB are assumed, similar to the specifications in Table I. The exact description of the simulation setup can be found in [14]. Concretely, Fig. 5b



(a) (b)

Fig. 5: (a) Nonlinear adaptive self-interference canceller structure for a 2×2 full-duplex MIMO device and (b) Signal-to-interference-plus-noise ratio (SINR) at canceller output as a function of transmit power in a 2×2 MIMO full-duplex device. I/Q modulator mismatch and power amplifier nonlinearity are present. SINR levels with linear, widely-linear, and two different nonlinear cancellers are shown. Results from [14] © IEEE.

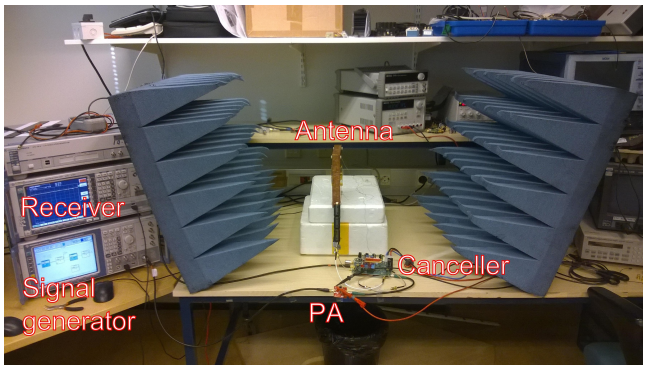
shows the signal-to-interference-plus-noise ratio (SINR) after digital cancellation for different methods, such as basic linear cancellation [16], nonlinear cancellation with no I/Q imbalance [12], widely-linear cancellation with no PA nonlinearity [13], and nonlinear cancellation modeling all previous impairments [14], referred to as joint cancellation in the figure. Strictly speaking, the number of basis functions of the joint nonlinear canceller deployed in [14] is 12 times larger than that of the plain linear cancellation scheme, but as shown in [14], only six of these are essential for the implementation. Furthermore, as shorter filters can be used for memory modeling with nonlinear terms, compared to the linear component, the total number of essential parameters to be estimated is still feasible. The nonlinear canceller from [14] is able to obtain SINR results within 1 dB of the optimal case when the transmit power remains below $+23$ dBm. The widely-linear canceller from [13] performs almost optimally when the transmit power is under $+15$ dBm, while nonlinear cancellation from [12] and plain linear cancellation fail to achieve good performance because of their inability to compensate the I/Q mismatch, which is the major source of interference in this scenario. Actual RF measurement examples with the parallel Hammerstein based nonlinear digital canceller in the single antenna case will be presented in the next section.

From the simulations results, we can conclude that digital cancellation techniques in combination with antenna isolation and RF cancellation can almost entirely remove the self-interference from the full-duplex device, whenever the cancellation architecture is able to accurately model the SI channel. A nonlinear cancellation architecture significantly improves the mitigation levels provided by its linear counterpart when operating with imperfect RF components, while adaptive schemes are able to track variations in the device environment that lead to changes within the SI channel.

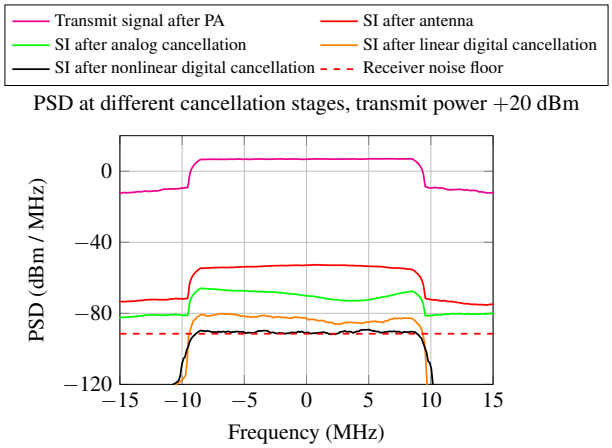
V. Overall Demonstrator System and Performance Evaluations

In order to evaluate the overall SI cancellation performance of a relay transceiver, i.e., a relay-antenna with analog/RF and digital SI cancellation, a measurement setup depicted in Fig. 6a is constructed. It consists of a vector signal generator, a PA, the relay antenna, an analog RF canceller, and a spectrum analyzer acting as an I/Q receiver. Digital cancellation processing stages, linear and nonlinear, are implemented on host computer processor. In the measurements, the aforementioned relay antenna is used, but without the wavetraps, to test the performance limits of the active analog/RF and digital cancellers. The RF center frequency in the measurements is 2.47 GHz.

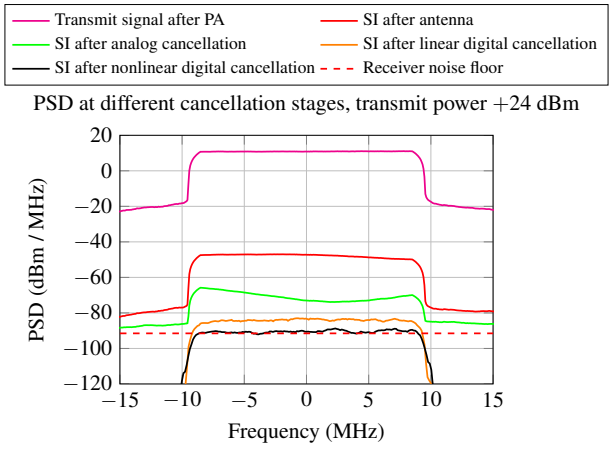
The analog/RF canceller is a novel design based on a paper by Y.-S. Choi and H. Shirani-Mehr [19]. The canceller consists of two cancellation taps delayed such that the delay of the main SI component lies between the aforementioned canceller taps. The reference signal for these taps is taken from the PA output using a directional coupler. The phase and amplitude of these reference signals are adjusted manually using vector modulators such that a good cancellation level is observed from the spectrum analyzer, indicating that the cancellation signal matches well with the actual SI signal. Cancellation is achieved simply by subtracting the cancellation signal from RX signal.



(a)



(b)



(c)

Fig. 6: (a) Measurement setup and equipment, (b) transmit signal and SI signal spectra at the different cancellation stages for 20 dBm (Texas Instruments CC2595 power amplifier) and (c) 24 dBm (MiniCircuits ZVE-8G power amplifier) transmit powers. The RF center frequency in the measurements is 2.47 GHz.

The rest of the measurement equipment are standard laboratory equipment, except for the PA, Texas Instruments CC2595, which is a low-cost chip intended for commercial use. It is deliberately chosen in order to demonstrate the performance of the digital nonlinear DSP algorithms under practical circumstances, where the PA distorts the transmit signal heavily. As already discussed, this is inevitable in many applications due to restrictions on the size, cost, and power consumption of the PA. Furthermore, also another PA, MiniCircuits ZVE-8G, is deployed to demonstrate the operation with higher output power and enhanced linearity.

Figures 6b and 6c depict then the overall measurement results obtained using the setup of Fig. 6a. The test signal is configured to be a 20 MHz wide OFDM signal at 2.47 GHz center frequency with an average transmit power of $+20$ dBm (TI PA case) or $+24$ dBm (MiniCircuits PA case), respectively, to test challenging scenarios with fairly high power levels. In Figs. 6b and 6c, the uppermost spectrum is that of the actual transmit signal, measured at the output of the PA. The high distortion levels at the PA output are evident also from the significant spectral regrowth visible outside the actual signal band. This is particularly clear in Fig. 6b where the used PA is operating already in highly nonlinear region. In Fig. 6c, in turn, the deployed PA is somewhat more linear, but clear nonlinear distortion is created also in this case.

After the PA output, the transmit signal propagates to the receiving antenna, turning into self-interference. From Figs. 6b and 6c it can be observed that the SI signal is attenuated by approximately 60 dB, regardless of the transmit power, when propagating to the receiver, including also 4 dB of losses from the setup. This result is well in line with the observations made in Fig. 4c, where SI attenuations of approximately 55 dB were measured for the antenna without the wavetraps.

The following trace in the figure is the residual SI after analog RF cancellation, which was performed using the aforementioned RF canceller prototype. Depending on the transmit power, the RF canceller is able to attenuate the SI signal by 15 or 22 dB. After this, the data is recorded to digital I and Q samples for offline post-processing with the nonlinear digital cancellation algorithm. In the digital canceller, nonlinear basis functions up to 7th order are considered, each of them having 10 pre-cursor and 10 post-cursor coefficients. The actual parameter estimation is carried out using block-wise least-squares, as discussed earlier. From Figs. 6b and 6c it can be observed that the residual SI power after digital nonlinear cancellation is at the IQ receiver noise floor (-91.5 dBm/MHz) with both transmit powers, meaning that the effective interference-plus-noise floor is suppressed by as much as 98 dB in total. This indicates that the total amount of SI cancellation from the antenna, RF canceller, and digital canceller is sufficient to attenuate the SI signal below the receiver noise floor, even with high transmit powers. Also note that, with ordinary linear digital cancellation, the power of the residual SI is significantly higher, evidencing that nonlinear modeling is indeed required in the digital cancellation processing of a practical in-band full-duplex transceiver. This is particularly clear in Fig. 6b where the PA is operating in highly nonlinear region.

VI. Conclusions

In this article, we explored sophisticated techniques for self-interference mitigation and cancellation within multi-antenna in-band full-duplex relays of compact size suitable for local area type networks. A novel antenna design utilizing resonant wavetraps was reported and it was shown to provide substantially enhanced passive isolation, despite its compact size. We also presented novel nonlinear and adaptive digital cancellation algorithms, which can enable enhanced digital self-interference cancellation levels when operating under practical nonlinear RF components. Furthermore, by combining the advanced antenna design with active RF and proposed digital canceller, we showed with actual RF measurements that the self-interference can be suppressed below the receiver noise floor, even when using regular low-cost components. The total measured aggregate self-interference suppression obtained in the measurements was close to 100 dB when operating at the transmit power level of $+24$ dBm. Thus, dealing with the self-interference problem in in-band full-duplex relays was shown to be technologically feasible, and hence it has the potential to significantly improve the performance in, e.g., local area type networks.

Acknowledgments

The authors wish to thank Prof. Sergei Tretyakov and Dr. Jari Holopainen, Aalto University, and Mr. Petteri Liikkanen, Tampere University of Technology, for technical guidance and assistance leading to these results. We also wish to thank Dr. Yang-seok Choi, Intel Corporation, for extensive technical consultation and advice related to active RF canceller prototype implementation.

References

[1] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[2] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.

[3] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. SIGCOMM'13*, Aug. 2013, pp. 375–386.

[4] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[5] H. Kang and S. Lim, "High isolation transmitter and receiver antennas using high-impedance surfaces for repeater applications," *Journal of Electromagnetic Waves and Applications*, vol. 27, no. 18, pp. 2281–2287, Sep. 2013.

[6] D. Korpi, S. Venkatasubramanian, T. Riihonen, L. Anttila, S. Otewa, C. Icheln, K. Haneda, S. Tretyakov, M. Valkama, and R. Wichman, "Advanced self-interference cancellation and multiantenna techniques for full-duplex radios," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013, pp. 3–8.

[7] P. Lindberg and E. Ojefors, "A bandwidth enhancement technique for mobile handset antennas using wavetraps," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 8, pp. 2226–2233, Aug. 2006.

[8] M. Karaboikis, C. Soras, G. Tsachtsiris, and V. Makios, "Compact dual-printed inverted-F antenna diversity systems for portable wireless devices," *IEEE Antennas and Wireless Propagation Letters*, vol. 3, no. 1, pp. 9–14, Dec. 2004.

[9] S. N. Venkatasubramanian, L. Li, C. Icheln, F. Ferrero, C. Luxey, and K. Haneda, "Impact of neutralization on isolation in co-planar and back-to-back antennas," in *Proc. 9th European Conference on Antennas and Propagation (EuCAP)*, Apr. 2015, pp. 1–5.

[10] S. N. Venkatasubramanian, A. Lehtovuori, C. Icheln, and K. Haneda, "On the constrains to isolation improvement in multi-antenna systems," in *Proc. 9th European Conference on Antennas and Propagation (EuCAP)*, Apr. 2015, pp. 1–5.

[11] M. E. Knox, "Single antenna full duplex communications using a common carrier," in *Proc. IEEE 13th Annual Wireless and Microwave Technology Conference*, Apr. 2012, pp. 1–6.

[12] L. Anttila, D. Korpi, V. Syrjälä, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *Proc. 47th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov. 2013, pp. 1193–1198.

[13] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.

[14] L. Anttila, D. Korpi, E. Antonio-Rodríguez, R. Wichman, and M. Valkama, "Modeling and efficient cancellation of nonlinear self-interference in MIMO full-duplex transceivers," in *Proc. IEEE Globecom Workshops*, Dec. 2014, pp. 862–868.

[15] V. Syrjälä, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 2977–2990, Jun. 2014.

[16] E. Antonio-Rodríguez, R. López-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "Adaptive self-interference cancellation in wideband full-duplex decode-and-forward MIMO relays," in *Proc. IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2013, pp. 370–374.

[17] ——, "SINR optimization in wideband full-duplex MIMO relays under limited dynamic range," in *Proc. 8th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jun. 2014, pp. 177–180.

[18] A. S. Tehrani, H. Cao, S. Afsardoost, T. Eriksson, M. Isaksson, and C. Fager, "A comparative analysis of the complexity/accuracy tradeoff in power amplifier behavioral models," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, no. 6, pp. 1510–1520, Jun. 2010.

[19] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.

Mikko Heino [S'14] received the M.Sc.(tech.) degree from Aalto University School of Electrical Engineering, Espoo, Finland, in 2014, where he is currently pursuing the D.Sc. degree in radio engineering at the Department of Radio Science and Engineering. His current research interests include antenna isolation improvement methods and antenna design towards in-band full-duplex systems. In 2014, Mikko received the second prize in the 2014 IEEE AP-S Student Design Contest as a part of the Aalto ELEC team.

Dani Korpi [S'14] received the B.Sc. and M.Sc. degrees (both with honors) in communications engineering from Tampere University of Technology (TUT), Finland, in 2012 and 2014, respectively. He is currently a researcher at the Department of Electronics and Communications Engineering at TUT, pursuing the D.Sc. (Tech.) degree in communications engineering. His main research interest is the study and development of single-channel full-duplex radios, with a focus on analyzing the RF impairments.

Timo Huusari received his B.Sc. Degree (with honors) in electrical engineering from Tampere University of Technology (TUT), Finland, in 2012. He is currently a research assistant at the Department of Electronics and Communications Engineering at TUT, pursuing the M.Sc. (Tech.) degree in electrical engineering. His main research interest is the study and development of full-duplex radios, with focus on self-interference cancellation device design and implementation.

Emilio Antonio-Rodríguez received the B.Sc. and M.Sc. degrees in Telecommunication Engineering, and Signal Processing Applications for Communications, from the University of Vigo, Spain, in 2006 and 2010, respectively. From 2008 to 2012, he worked as a researcher at the Galician Research and Development Center in Advanced Telecommunications (GRADIANT). Since 2012, he is pursuing the D.Sc. (Tech.) degree at the Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, Finland. His research interests are related to adaptive signal processing, interference suppression and signal processing for communications.

Sathya N. Venkatasubramanian received his B.E degree from Anna University, Chennai, India in 2010 and received his M.Sc.(tech.) degree in January 2013 majoring in radio science and engineering from Aalto University School of Electrical Engineering, Espoo, Finland. Since 2013 he is a doctoral candidate in the Department of Radio Science and Engineering at Aalto University School of Electrical Engineering. His current research interests include antennas and propagation towards in-band full-duplex systems.

Taneli Riihonen received the D.Sc. degree in electrical engineering (with distinction) from Aalto University, Helsinki, Finland in August 2014. Since September 2005, he has held various research positions at the Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering. He is currently a visiting scientist at Columbia University in the City of New York, USA, and serving in the editorial board of IEEE Communications Letters. His research activity is focused on physical-layer OFDM(A), multiantenna, relaying and full-duplex wireless techniques.

Lauri Anttila [S06, M11] received the M.Sc. degree and D.Sc. (Tech) degree (with honors) from Tampere University of Technology (TUT), Tampere, Finland, in 2004 and 2011. Currently, he is a postdoctoral research fellow at the Department of Electronics and Communications Engineering at TUT. His research interests include signal processing for wireless communications, transmitter and receiver linearization techniques, and full-duplex radio systems.

Clemens Icheln [S06, M11] received the M.Sc. degree in electrical engineering (Dipl.-Ing.) at Hamburg-Harburg University of Technology, Germany, in 1996, the Licentiate degree in radio engineering and the Doctor of Science in Technology degree at Helsinki University of Technology, Espoo, Finland, in 1999 and 2001, respectively. He is currently working as University Lecturer at the Department of Radio Science and Engineering of Aalto University School of Electrical Engineering. His main research interests are the design of antennas for mobile terminals and for compact relays, and their evaluation methods.

Katsuyuki Haneda [S03, M07] received the Doctor of Engineering degree from the Tokyo Institute of Technology, Tokyo, Japan, in 2007. Dr. Haneda is presently an assistant professor in the Aalto University School of Electrical Engineering. His current research activity focuses on high-frequency radios such as millimeter-wave and beyond, wireless for medical and post-disaster scenarios, radio wave propagation prediction, and in-band full-duplex radio technologies.

Risto Wichman received his Ph.D. from Tampere University of Technology, Finland in 1995. From 1995 to 2001, he worked at Nokia Research Center, Finland, as a senior research engineer. In 2002, he joined Department of Signal Processing and Acoustics, Aalto University School of Electrical Engineering, Finland, where he is a full professor since 2008. His research interests include signal processing techniques for wireless communication systems.

Mikko Valkama [S'00, M'02] received his M.Sc. and Ph.D. Degrees (both with honors) from Tampere University of Technology (TUT), Finland, in 2000 and 2001, respectively. In 2003, he was working as a visiting researcher with the Communications Systems and Signal Processing Institute at SDSU, San Diego, CA. Currently, he is a Full Professor and Department Vice-Head at the Department of Electronics and Communications Engineering at TUT, Finland. His general research interests include communications signal processing, cognitive radio, full-duplex radio, radio localization, and 5G mobile cellular radio systems.

Publication 4

D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: Pushing the limits," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 80–87, Sep. 2016. DOI: 10.1109/MCOM.2016.7565192

Full-Duplex Mobile Device – Pushing the Limits

Dani Korpi, Joose Tamminen, Matias Turunen, Timo Huusari,
Yang-Seok Choi, Lauri Anttila, Shilpa Talwar, and Mikko Valkama

***Abstract*—In this article, we address the challenges of transmitter-receiver isolation in *mobile full-duplex devices*, building on shared-antenna based transceiver architecture. Firstly, self-adaptive analog RF cancellation circuitry is required, since the capability to track time-varying self-interference coupling characteristics is of utmost importance in mobile devices. In addition, novel adaptive nonlinear DSP methods are also required for final self-interference suppression at digital baseband, since mobile-scale devices typically operate under highly nonlinear low-cost RF components.**

In addition to describing above kind of advanced circuit and signal processing solutions, comprehensive RF measurement results from a complete demonstrator implementation are also provided, evidencing beyond 40 dB of active RF cancellation over an 80 MHz waveform bandwidth with a highly nonlinear transmitter power amplifier. Measured examples also demonstrate the good self-healing characteristics of the developed control loop against fast changes in the coupling channel. Furthermore, when complemented with nonlinear digital cancellation processing, the residual self-interference level is pushed down to the noise floor of the demonstration system, despite the harsh nonlinear nature of the self-interference. These findings indicate that deploying the full-duplex principle can indeed be feasible also in mobile devices, and thus be one potential technology in, e.g., 5G and beyond radio systems.

I. Introduction

INBAND full-duplex communications is widely regarded as one potential solution towards more spectrally efficient wireless networks. The basic idea behind it is to utilize the available temporal and spectral resources to the fullest extent by transmitting and receiving data signals simultaneously at the same center frequency [1], [2]. In theory, this will result in doubling of the radio link data rate while requiring no additional bandwidth. Furthermore, when combined with proper scheduling in multiuser networks, this can be translated into an increase also in the cell and network capacities [3]. Especially in the future 5G era, inband full-duplex communications can be one enabler and crucial step towards the desired 1000-fold increase in the total throughput [4], [5]. Thus, implementing a fully functional inband full-duplex transceiver is a tempting prospect.

However, in practice, realizing the potential performance gains is far from trivial, as extremely efficient attenuation of the own transmit signal is required. Note that now it is not possible to filter out the own transmission with, e.g., a duplexer, since it is overlapping with the actual received signal of interest in the frequency domain. In theory, canceling this so-called self-interference (SI) can be done by subtracting the own transmit signal from the total received waveform. In practice, on the other hand, the SI signal will always be distorted in a linear as well as nonlinear manner while propagating to the receiver, and thereby it is not a trivial task to reproduce a sufficiently accurate cancellation signal. Attenuating the SI signal by an adequate amount is in fact the central research problem, which must be resolved in order to implement a practical inband full-duplex radio [1], [2], [5], [6].

The nonlinear distortion due to analog impairments is an especially prevalent issue in mobile-scale devices, which typically utilize low-cost mass-produced RF components. For this reason, a typical assumption in the reported works has been that, in a mobile cellular network, the base station (BS) is able to communicate in full-duplex mode, whereas the mobiles are legacy half-duplex devices [7]. An illustration of this type of a solution is shown in Fig. 1(a), where a full-duplex capable BS is serving half-duplex mobile users. The benefit of this solution is that it avoids the challenges of implementing a mobile full-duplex transceiver, and instead requires only the BS to be full-duplex capable. This is a significantly easier prospect, since the BS typically utilizes more expensive higher quality components, and it can also have a significant amount of spatial isolation between the transmitter and the receiver.

However, limiting the full-duplex operation only to the BS side does not obviously capitalize the full potential of the full-duplex principle. By having also the mobile devices full-duplex capable, the overall data rate of the corresponding cell could be significantly increased [8]. This is illustrated in Fig. 1(b), where the BS can now exchange data in a full-duplex manner with each mobile user, resulting in improved spectral efficiency and higher data rates. For this reason, in this article we will *investigate the possibilities and challenges of implementing full-duplex capable mobile devices*. Due to extreme size and cost constraints in mobile devices, as well as stringent requirements regarding the power consumption, this is an extraordinarily challenging task [9].

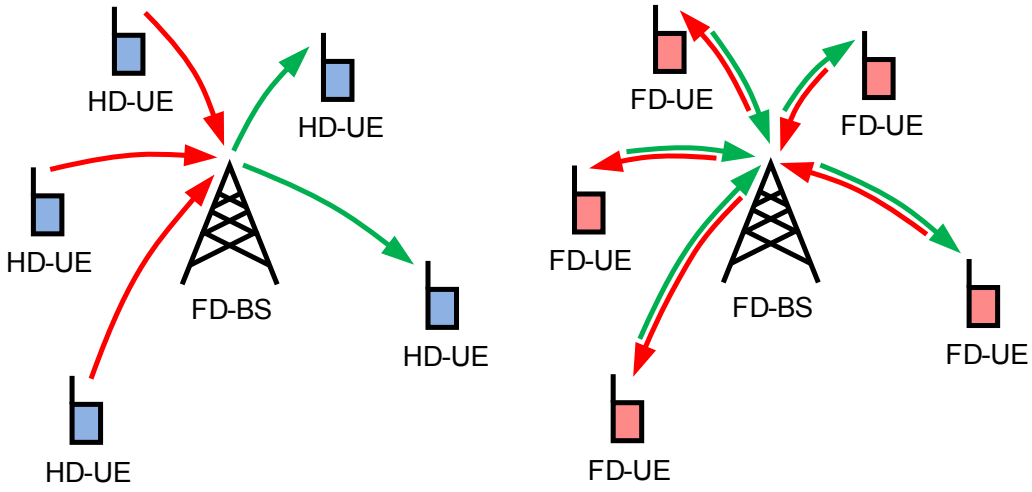
D. Korpi, J. Tamminen, M. Turunen, L. Anttila, and M. Valkama, are with the Department of Electronics and Communications Engineering, Tampere University of Technology, PO Box 692, 33101, Tampere, Finland, e-mail: dani.korpi@tut.fi, joose.tamminen@tut.fi, matias.turunen@tut.fi, lauri.anttila@tut.fi, mikko.e.valkama@tut.fi.

T. Huusari is with Intel Corporation, Tampere, Finland, e-mail: timo.s.huusari@intel.com.

Y.-S. Choi is with Intel Corporation, Hillsboro, OR 97124, USA, e-mail: yang-seok.choi@intel.com.

S. Talwar is with Intel Corporation, Santa Clara, CA 95054-1549, USA, e-mail: shilpa.talwar@intel.com.

The research work leading to these results was funded by the Academy of Finland (under the project #259915 "In-band Full-Duplex MIMO Transmission: A Breakthrough to High-Speed Low-Latency Mobile Networks"), Nokia Foundation, Tuula and Yrjö Neuvo Research Fund, and Emil Aaltonen Foundation.



(a) (b)

Fig. 1: (a) An illustration of a single cell where the base station (BS) is full-duplex capable and the mobiles (UEs) are legacy half-duplex devices, and (b) a similar illustration of a single cell where all parties are full-duplex capable.

Firstly, due to the restricted dimensions in a mobile-scale full-duplex transceiver, there is likely to be no space for separate transmit and receive antennas. This means that the transmitter and receiver must share an antenna, while still maintaining a reasonable amount of isolation between each other. There have already been some preliminary demonstrations where this has been implemented in practice, and thereby this aspect of a mobile inband full-duplex transceiver is potentially feasible [6], [9].

In addition, because of the wide bandwidths of the modern radio systems, advanced wideband cancellation processing in the analog/RF domain is required also in full-duplex mobile devices. The feasibility of this type of wideband RF cancellation circuits, utilizing, e.g., several delay lines and appropriate amplitude and phase tuning to model and track the frequency and time dependencies of the wideband SI channel, have also been preliminarily demonstrated in practice [6], [10]. The remaining challenge is implementing a wideband RF canceller in mobile scale, such that the transmitter and receiver utilize the same antenna.

Another aspect, which becomes a considerable factor and concern in mobile devices, is the quality of the analog components. Specifically, the low-cost components, which are typically used in mass-produced handheld devices, distort the SI signal such that linear digital processing alone cannot reproduce and cancel the residual SI waveform accurately enough [6], [9]. This means that advanced modeling and processing, taking into account the different analog impairments, is required in order to produce a sufficiently accurate cancellation signal. For instance, modeling nonlinear distortion in the digital SI regeneration and cancellation stage has been shown to improve the performance of a practical inband full-duplex transceiver [6], [9].

In this article we will take a closer look into the aforementioned challenges. In addition, we will also present some of our recent findings for solving them and demonstrate with an actual prototype implementation that the challenges caused by the limited size and RF component quality of mobile-scale devices can be tackled by incorporating state-of-the-art algorithms and cancellation processing. The obtained measurement results show that the prototype is in fact able to cancel the self-interference to the level of the receiver noise floor. In particular, this article builds partially on the recent scientific findings by the authors, reported primarily in [9]–[12].

II. Shared-Antenna Mobile Full-Duplex Device Architecture

The general structure of a shared-antenna mobile-scale direct-conversion full-duplex transceiver is shown in Fig. 2(a), which illustrates all the relevant aspects required to achieve real full-duplex operation under practical conditions. The key component enabling the transmitter and receiver to share a single antenna in a mobile full-duplex device is a circulator, which is used to connect the antenna to the transceiver. A circulator is a three port device which steers the signal through its ports such that it comes in at one port and then exits the circulator from the next port, depending on the direction of rotation (i.e., clockwise or counterclockwise). In principle, the signal cannot propagate to the opposite direction, which ensures a certain amount of isolation between the transmitter and the receiver. Depending on the size and cost of the circulator, typical practical values for the isolation vary between 20 and 60 dB, while the attenuation in the desired direction is usually less than half a decibel. Being a passive component, its size is ultimately dictated by the wavelength of the operating frequency.

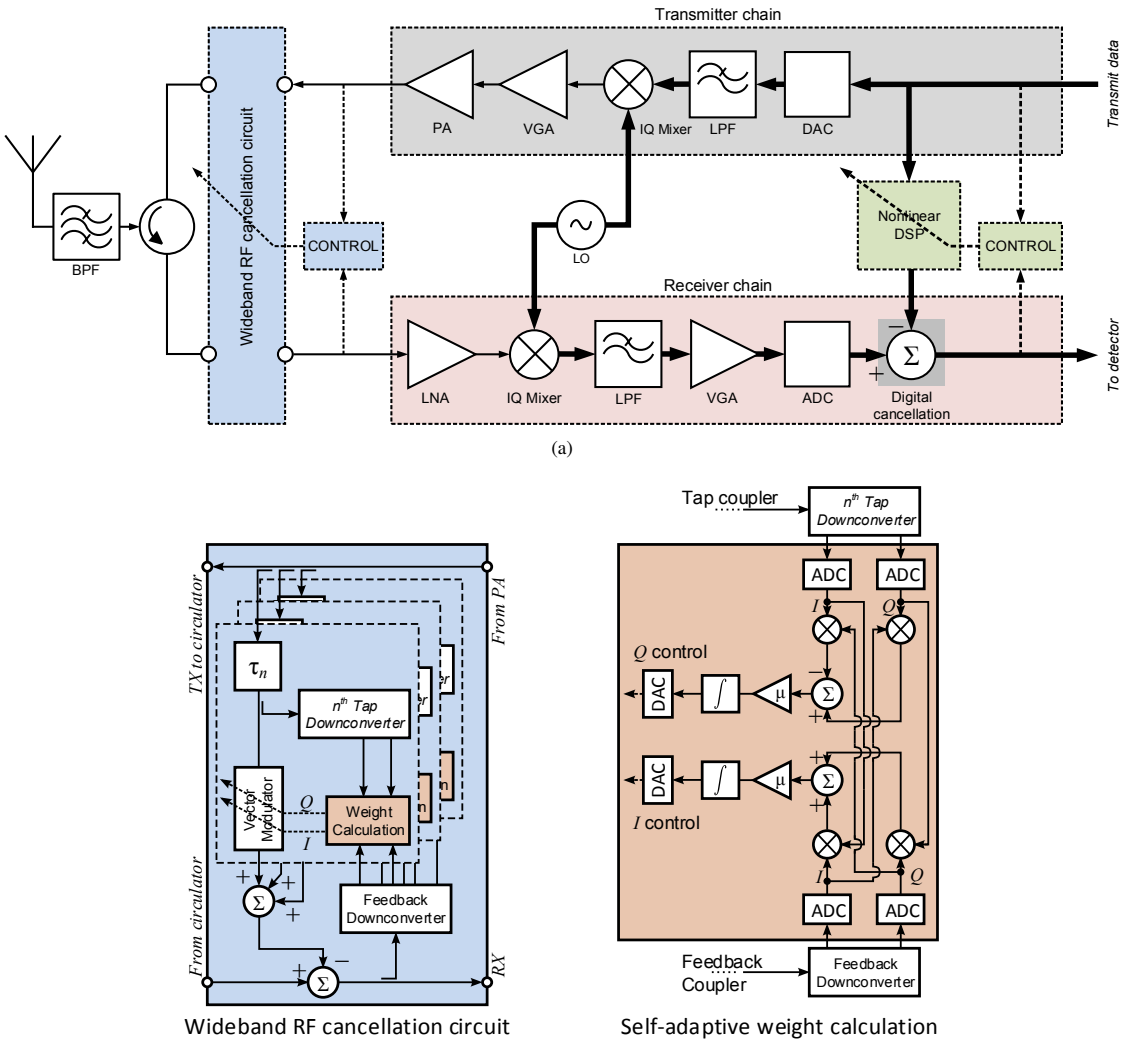


Fig. 2: (a) The considered shared-antenna mobile full-duplex transceiver architecture, including self-adaptive RF cancellation and self-adaptive nonlinear digital cancellation, alongside with (b) a general block diagram depicting the RF canceller and the self-adaptive weight calculation for one tap.

Another option for isolating the transmitter and receiver sharing the same antenna is to use an electrical balance duplexer, which essentially tries to mimic the impedance of the antenna with a tunable balance network [13]. With the help of a hybrid transformer, this results in a significant amount of isolation between the transmitter and the receiver. An electrical balance duplexer can be implemented in a more compact form than a circulator, which makes it very suitable for mobile devices, but it also suffers from some inherent insertion loss. Another drawback of the electrical balance duplexer is the need to actively tune the impedance network, since the impedance of the antenna is time-variant. This is obviously not needed in a circulator-based system. For these reasons, in this article we focus on a demonstrator implementation utilizing a circulator, since it is a passive device and thereby more robust. Nevertheless, for future implementations, the electrical balance duplexer is also a very prominent candidate.

When using the circulator-based architecture, there are two strong components in the SI signal, observed towards the receiver path. Firstly, there will be leakage through the circulator, whose magnitude can be estimated by subtracting the amount of circulator isolation from the transmit power. Here, the SI is usually attenuated by at least 20 dB, as already mentioned. The second strong component is the power reflected

by the mobile antenna, caused by the impedance mismatch at its input. If the input was perfectly matched, the antenna would accept all the supplied power and this SI component would not exist, but in practice the mismatch will always cause a part of the power to be reflected back to the transmission line. When using off-the-shelf antennas, matching values better than -20 dB are seldom obtained, and the reflection from the antenna constitutes a significant portion of the total SI, potentially even dominating compared to direct leakage through the circulator.

Weaker components in the composite SI signal come mainly from the multipath reflections, which propagate back to the antenna from the surrounding environment. They are heavily dependent on the type of environment around the antenna but usually the multipath reflections will be significantly weaker than the leakage through the circulator or the reflection from the antenna. However, a change in the near field of the antenna (e.g., wrapping a hand around it) affects its matching, which will directly change the amount of reflected power.

Because of the leakage through the circulator, as well as the reflections coming from the antenna and the surrounding environment, additional SI attenuation is typically required, both in the analog/RF and digital domains. In general, the overall analog attenuation of the SI signal prior to entering the receiver chain must be sufficient to ensure that

- The SI power level is not too high for the receiver low-noise amplifier (LNA), to prevent receiver saturation
- The dynamic range of the analog-to-digital-converters (ADCs) is high enough to capture the residual SI as well as the weak received signal of interest with sufficient precision.

Depending on the receiver, either of these can be the limiting factor [11]. Usually, the passive SI attenuation provided by circulator isolation and antenna matching is clearly insufficient to ensure these requirements [10], [11]. This creates a strong motivation for active RF cancellation, which provides additional SI suppression before the actual receiver chain by subtracting a modified copy of the transmit signal from the overall received signal.

Remembering again the wide bandwidth of the signals that are used in modern cellular networks, it is obvious that the active RF canceller within a mobile full-duplex device must be capable of efficient *wideband cancellation*. This can be ensured by having a multi-tap analog SI canceller where several differently delayed copies of the transmit signal are used as reference signals, each of them having tunable amplitude and phase. The objective of the RF cancellation circuit is then to control the phases and amplitudes of these reference signals such that the SI signal is suppressed. This type of an RF canceller is capable of modeling the coupling channel over significantly wider bandwidths than the conventional solutions [6], [10].

Another important consideration for a mobile full-duplex device is to also have sufficient adaptivity in the RF canceller, a factor that has thus far been neglected in most of the reported works [2]. This is needed to support efficient cancellation of the SI under a time-varying channel environment, caused by the moving objects in the vicinity of the device, such as a person walking by, or by the movement of the device itself. The tracking of the SI channel can be done by using either digital or analog circuits to control the phases and amplitudes of the cancellation signals in a *self-adaptive* manner [10]. This topic, alongside with other aspects related to the RF canceller, is elaborated in more details in Section III.

Furthermore, due to the high power level of the received SI signal, analog SI cancellation alone is typically not enough to attenuate it below the receiver noise floor. Thus, the final attenuation of the residual SI must be done in the *digital domain*. There, the cancellation signal can be constructed from the original transmit data by filtering it in accordance with the remaining effective SI channel. One important benefit of digital SI cancellation is the relatively easy inclusion of nonlinear modeling of the SI waveform, which can be done conveniently by utilizing *nonlinear basis functions* [6], [9], as well as the natural support for self-tracking of the SI channel characteristics through adaptive filtering. As has been demonstrated recently, nonlinear modeling can significantly improve the SI cancellation performance in a practical full-duplex transceiver [6], [9]. Thus, nonlinear adaptive digital signal processing is also a key feature in a mobile full-duplex device to ensure efficient cancellation and tracking of residual SI. This is addressed in more details in Section IV.

III. Advanced Self-Adaptive RF Cancellation Principle

This section describes the detailed principle of such an RF cancellation circuit that fulfills the aforementioned requirements regarding wideband operation and self-adaptivity [10]. A prototype implementation of this type of an RF canceller is then reported and measured in Section V. As outlined above, the RF canceller aims at reconstructing and canceling the received composite SI waveform, which consists of various components with different delays. Since the delays of these components are unknown and also time varying, the delays in the SI regeneration and cancellation paths are not equal to the true delays. Hence, as a whole, the active RF canceller can be seen as an interpolator which tries to regenerate and track the true composite SI, by using *predefined delays but tunable amplitudes and phases* [10].

Overall, optimum linear filtering for interference or echo cancellation is a thoroughly studied field in the literature. However, most of the reported research and implementations focus on digital baseband while here our focus is fully at analog RF domain. In the full-duplex radio field, optimum filtering based analog RF cancellation has been recently addressed, in terms of passband analog finite impulse response (FIR) filtering, in [6] and [10]. The filter coefficients or weights can be obtained in analog [10] or digital domain [6], and in adaptive [10] or non-adaptive [6] manner. Unlike in [6], *per-tap phase shifting* is also included in [10], allowing phase rotation of a tapped delayed signal at passband. Thus, the filter tap weights in [10] become complex, when interpreted from the baseband waveform perspective, while those in [6] are real. These effectively complex taps significantly reduce the cancellation performance dependency on frequency, tap delays,

and the underlying true delays of the different SI components, thereby also reducing the number of taps required. This, in turn, is crucial for mobile devices in order to minimize cost, size, and power consumption.

In general, there are two options for obtaining and controlling the tap weights: an *open loop* and a *closed loop*. In the *open loop*, separate SI channel estimation is needed, followed by the actual canceller weight calculations. Such strategy obviously calls for digital processing of a large amount of data, thereby producing a significant delay in the canceller adaptation. In the *closed loop*, on the other hand, the weights are directly optimized to minimize the SI power at the canceller output. Such closed-loop adaptive processing structure is thus essentially a negative feedback system, where the weights are automatically adjusted in real-time to keep the residual SI power low at the canceller output. This strategy is very well suitable for directly tracking a time-varying SI channel under strict delay requirements [10].

For this reason, in a mobile device, the weight adaptation for RF cancellation must be done in a closed-loop fashion since tracking the characteristics of the overall SI waveform in real-time is a crucial feature. The general structure for such a closed-loop wideband RF canceller circuit utilizing three taps is illustrated in Fig. 2(b), where the LMS-based learning algorithm, operating in the digital domain, is also shown for a single tap. This type of a canceller structure has been observed to provide excellent cancellation performance under a wide bandwidth and highly varying channel conditions, as will be shown through measurements in Section V. Furthermore, as discussed in detail in [10], this type of an RF canceller is very robust against various circuit imperfections that typically occur in mobile-scale devices. In particular, deploying the power amplifier (PA) output as the reference signal in RF cancellation is beneficial, since this way all the main transmit chain imperfections are automatically included in the cancellation signal [10], [11], and hence subtracted along with the linear SI. This is particularly important when utilizing a mobile-scale power-efficient PA, which creates substantial nonlinear distortion. This then relaxes to certain extent the requirements on the final residual SI suppression at digital baseband, and also reduces the required dynamic range for the main receiver ADC.

IV. Adaptive Nonlinear Digital Cancellation for Final SI Suppression

After analog SI cancellation, the power level of the residual SI can still be relatively strong in the digitized signal. This calls for additional digital SI cancellation, which will then decrease the level of the SI signal below the receiver noise floor. The most straight-forward method for canceling SI in the digital domain is to use the original transmit data as the reference signal, which is then modified according to the effective channel experienced by the residual SI signal and subsequently subtracted from the overall received signal. The channel includes the effects of both the transmitter and the receiver, the circulator, and the RF canceller, as well as the multipath components reflected from the antenna and the surrounding environment. Modeling, estimating and tracking this effective SI channel is the key factor in digital SI cancellation, and it will determine the achievable cancellation in the digital domain.

Typically, in most works reported in the literature, the effective SI channel is assumed to be a linear multipath channel, which essentially means that the transmitter and receiver chains are assumed to be linear [2]. With high-quality components, e.g., well-calibrated laboratory equipment, this can indeed be the case. However, when considering a mobile-scale full-duplex transceiver utilizing low-cost mass-produced components, assuming the transmitter and receiver chains to be linear will result in a significant model mismatch. In particular, the transmitter PA is typically heavily nonlinear, especially with the higher transmit powers. This has a significant impact on the residual SI observed at digital baseband [9], [11].

Stemming from the above, the nonlinearity of the components must be considered in the digital cancellation processing, especially the nonlinear distortion produced by the transmitter PA [6], [9]. In principle, this can be done by modeling the residual SI as a weighted sum of nonlinear transformations of the original transmit data, each of which has also some delayed components (memory) present. A principal structure of such a nonlinear digital canceller is shown in Fig. 3, where the original transmit data is first transformed with nonlinear basis functions and then orthogonalized to ensure efficient learning, as discussed in more detail in [9]. In this nonlinear canceller, the actual transceiver chain is modeled as a cascade of a nonlinear PA and a linear filter, the latter of which consists of the PA memory, multipath components of the SI signal reflected from the surroundings, and the RF cancellation circuit [9]. This means that the nonlinear residual SI channel follows a parallel Hammerstein (PH) model, whose parameters are relatively straight-forward to estimate and track based on the observed SI signal [14].

In general, the parameter estimation can be carried out, for instance, with block least squares or least mean squares (LMS), depending on the application and available computational resources. In a practical mobile transceiver, adaptivity is a very important factor, as already discussed, and thus LMS or some other adaptive algorithm is preferred to ensure high performance under varying coupling channel conditions. In Figs. 2(a) and 3, the adaptivity is depicted by the real-time control block, which tunes the coefficients based on the canceller output signal. The digital canceller output signal is also used for the actual receiver digital baseband processing, including the detection of the actual received signal of interest.

Overall, the performance of this type of a nonlinear digital canceller is of course highly dependent on the validity of the underlying model. Namely, the utilized PH model has been shown to be quite accurate for modeling a wide variety of PAs [14], but often there are also other sources of impairments that are obviously not included in the model, such as phase noise [15]. However, under typical circumstances, the nonlinearity of the PA is the most significant analog impairment from the SI cancellation perspective [11], which means that the PH model can be expected to provide sufficient SI cancellation performance. Therefore, combining an adaptive nonlinear digital canceller with the proposed multi-tap adaptive RF

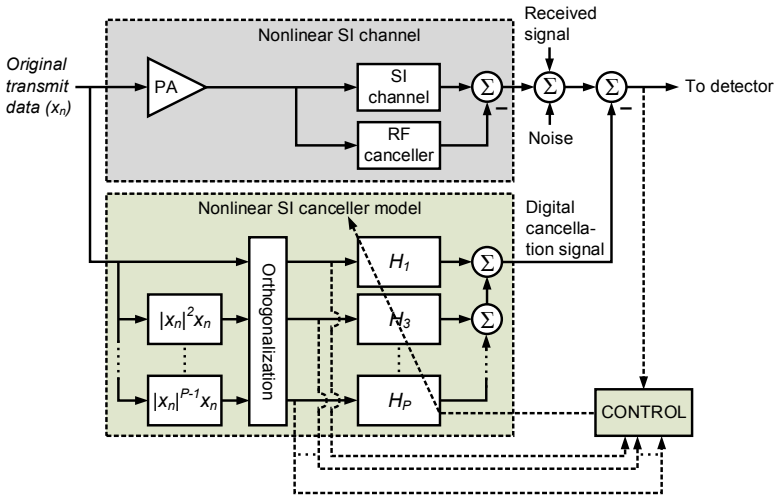


Fig. 3: A general illustration of adaptive nonlinear digital cancellation processing, incorporating behavioral models for a wideband nonlinear power amplifier and time-varying coupling characteristics.

canceller will result in an agile mobile full-duplex transceiver design that is capable of efficient SI cancellation both in the analog and digital domains.

V. Demonstrator Implementation and Measured Results

To evaluate the total performance of the described mobile full-duplex transceiver architecture, real-life RF measurements and experiments are performed with the measurement setup shown in Fig. 4, which integrates the different considered cancellation stages together. The measurements are carried out using a National Instruments PXIe-5645R vector signal transceiver (VST) both as a transmitter and a receiver, complemented with an external PA. The used transmit signal is an LTE waveform with an instantaneous bandwidth of 20, 40 or 80 MHz, centered at 2.46 GHz. The VST output is then connected directly to a Texas Instruments CC2595 PA which has a gain of 24 dB at the chosen input power level. The used PA is a commercial low-cost chip intended to be used in low-cost battery-powered devices. This means that the PA produces a significant amount of nonlinear distortion into the SI waveform, especially with the input power levels used in these measurements.

After the PA, the signal is divided between the RF canceller prototype interface and the antenna port using a directional coupler. Accounting for all the losses incurred by dividing the transmit signal among the different paths, the approximate transmit power at the antenna is in the order of +6. . .+8 dBm, depending on the bandwidth. Such transmit powers are feasible in future ultra-dense 5G networks where inter-site distances below 100 m are to be expected[1], while experimenting with higher transmit powers is an important topic for our future work. The deployed circulator and the low-cost shared-antenna yield an overall isolation only in the order of 20 dB between the transmitter and the receiver chains, mostly because of the reflection from the antenna. Then, after the circulator, the desired RX signal and SI are routed back to the prototype RF canceller, which performs the RF cancellation procedure utilizing the PA output signal as described in Section III. Finally, the RF cancelled signal is routed to the receiver (NI PXIe-5645R) and captured as digital I and Q samples, which are post-processed offline to implement linear as well as nonlinear digital baseband cancellation. The parameter learning in the digital canceller is done with basic LMS adaptation described in more detail in [9], using a highest nonlinearity order (P) of 11. Learning the parameters with such a simple algorithm guarantees that the digital cancellation procedure is done in a computationally efficient manner, which is obviously an important aspect in a mobile-scale device. In the forthcoming results, the adaptive digital cancellation algorithm is first allowed to converge towards the steady-state coefficient values, after which the cancellation performance is measured. This ensures that the results show the true performance of the digital canceller.

A. Self-Adaptive RF Canceller Implementation and Measured Performance

First, the performance of the self-adaptive three-tap RF canceller, illustrated in Fig. 2(b), is evaluated with different bandwidths. Figures 5(a), 5(b), and 5(c) show the measured RF signal spectra after the RF canceller when the signal bandwidth is 20, 40, and 80 MHz, respectively. Also the spectra of the transmit signal and the RF canceller input signal are shown for reference. With the 20 MHz signal, the overall suppression achieved by the RF canceller is close to 50 dB, and even with the 80 MHz signal the SI can still be attenuated by more than

[1]Nokia Solutions and Networks, "Ten key rules of 5G deployment," *white paper C401-01178-WP-201503-1-EN*, 2015

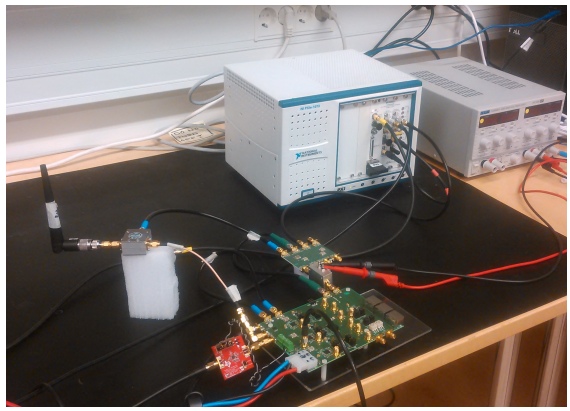
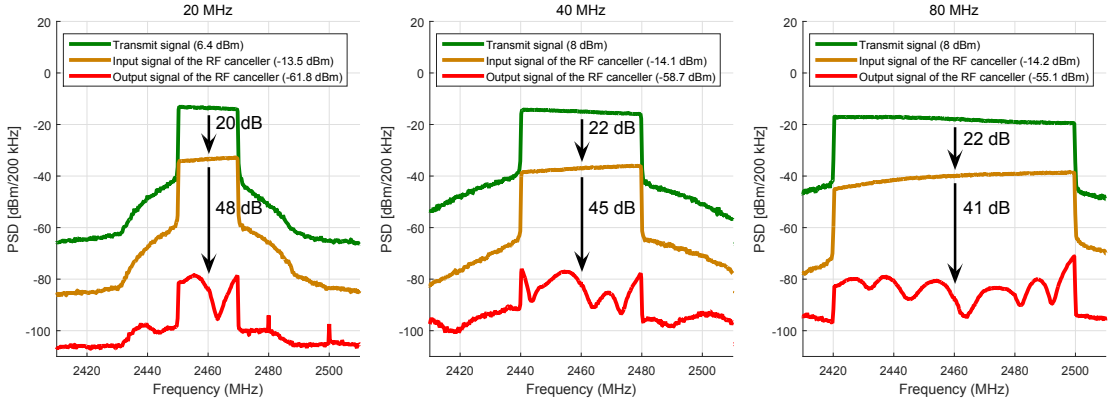


Fig. 4: The laboratory measurement setup used to evaluate and demonstrate the analog and digital self-interference cancellation performance of a shared-antenna full-duplex transceiver.



(a) (b) (c)

Fig. 5: The RF cancellation performance with a three-tap canceller at 2.46 GHz center-frequency with (a) 20 MHz, (b) 40 MHz, and (c) 80 MHz transmit signals. The value in the parentheses is the power measured over the mentioned bandwidth around the center frequency.

40 dB. To the best of our knowledge, these are the highest reported values for *active RF cancellation* thus far, especially for such wideband signals. By increasing the number of taps in the canceller, even wider bandwidths can be potentially supported with this architecture [10].

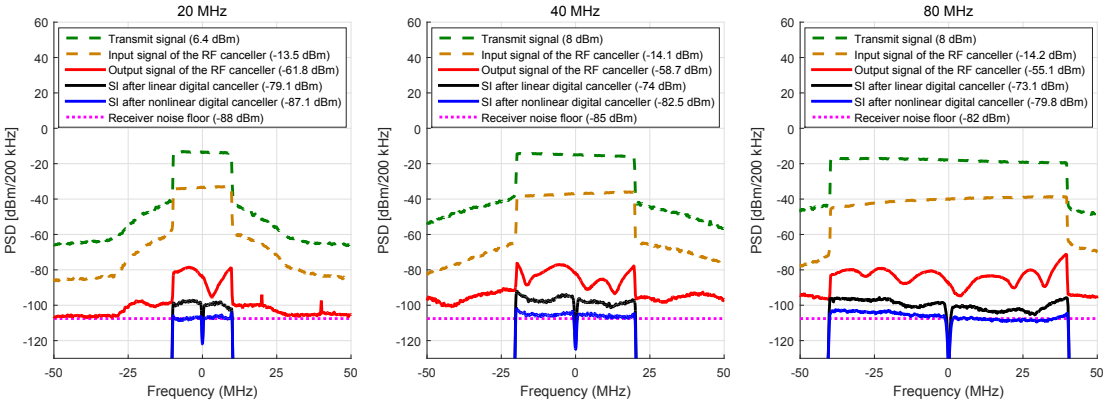
Next, we demonstrate the self-adaptation capabilities of the developed RF cancellation circuit and its underlying automated control intelligence. A time-varying reflection scenario around the antenna is deliberately created by bringing different reflecting materials close to the antenna. This changes heavily the total SI coupling channel, and hence calls for fast adaption in the RF cancellation circuit. This overall setup is demonstrated through a video recording of the experimentation, showing that the developed RF canceller can rapidly self-heal its operation by automatically tuning the amplitudes and phases of the RF cancellation paths. The video recording is available at

- RF canceller demonstration: http://ieeexplore.ieee.org or http://www.tut.fi/full-duplex/commag

B. Total Integrated System Performance Including Nonlinear Digital Cancellation

Despite the impressive RF cancellation performance demonstrated above, the level of the residual SI is still substantially above the receiver noise floor in the digital domain. Thus, the downconverted and digitized received signal is next processed by the digital SI canceller, which is implemented in a host processor. Furthermore, because of the low-cost mobile-scale PA, nonlinear digital SI cancellation with LMS parameter learning is utilized [9].

Figures 6(a), 6(b), and 6(c) show example signal spectra of the SI signal after each cancellation stage, including adaptive nonlinear digital cancellation described in Section IV. It can be observed that, for the considered bandwidths of 20, 40, and 80 MHz, the combined attenuation of the circulator and the RF canceller is 63–68 dB, after which the described adaptive nonlinear digital canceller attenuates the SI signal further by another 25 dB. With classical linear digital cancellation,



(a) (b) (c)

Fig. 6: The overall cancellation performance with (a) 20 MHz, (b) 40 MHz, and (c) 80 MHz LTE waveforms at 2.46 GHz, the contribution of each cancellation stage being shown separately (projected on to the baseband). In the digital canceller, nonlinearities up to the 11th order are considered to accurately model the low-cost power amplifier. This allows the nonlinear digital canceller to push the residual SI to the receiver noise floor, clearly outperforming the linear canceller whose performance is heavily limited by the power amplifier induced nonlinear distortion.

the total SI attenuation is roughly 10 dB less, illustrating that nonlinear digital cancellation processing is a necessary requirement for a mobile full-duplex transceiver with substantial nonlinear distortion in the transmitter power amplification stage.

Overall, these examples demonstrate that the described mobile full-duplex transceiver architecture is capable of attenuating the SI practically to the receiver noise floor even with the 80 MHz signal bandwidth. This indicates that very good overall wideband performance is achieved also under practical conditions, despite a heavily nonlinear transmitter PA and a shared antenna. Hence, the proposed architecture is well capable of coping with the challenges posed by the mobile device environment, such as circuit imperfections, wideband operation, and time-varying channel conditions.

VI. Conclusion

To fully capitalize the benefits of inband full-duplex radio technology in, e.g., cellular networks, also the mobile devices should support simultaneous transmission and reception at the same center-frequency. This article explored the most prominent challenges in implementing mobile inband full-duplex devices, and also described solutions to these problems. In particular, a mobile full-duplex device must be capable of shared-antenna operation, as well as adapting to the changes in the self-interference channel environment. Also, since a mobile-scale device is typically relying on low-cost components, the different circuit impairments must be considered as they directly affect the self-interference cancellation ability. Furthermore, nowadays also mobile devices must be able to handle very wideband signals to ensure high data rates.

As a solution to these challenges, this article demonstrated a prototype implementation having a shared transmit/receive antenna, an adaptive wideband multi-tap RF canceller, and an adaptive wideband nonlinear digital canceller. It was shown that the novel wideband RF cancellation solution with built-in capability to automatically track changes in the self-interference channel characteristics can yield RF cancellation gains beyond 40 dB, even with waveform bandwidths in the order of 80 MHz and a highly nonlinear low-cost power amplifier. Measured examples also demonstrated good self-adaptation capabilities against fast changes in the environment around the antenna. Moreover, the adaptive nonlinear digital canceller was shown to push the residual self-interference down to the noise floor of the receiver, also under a heavily nonlinear transmitter power amplifier. All in all, these findings pave the way towards potentially enabling the full-duplex capability also in the mobile devices of the future 5G or beyond radio communication systems.

References

[1] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[2] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[3] S. Goyal, P. Liu, S. S. Panwar, R. A. Difazio, R. Yang, and E. Bala, "Full duplex cellular systems: will doubling interference prevent doubling capacity?" *IEEE Communications Magazine*, vol. 53, no. 5, pp. 121–127, May 2015.

[4] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. Soong, and J. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[5] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114–121, Feb. 2014.

[6] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. SIGCOMM'13*, Aug. 2013, pp. 375–386.

[7] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *Proc. 45th Asilomar Conference on Signals, Systems and Computers*, Nov. 2011, pp. 2002–2006.

[8] A. Sahai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of full-duplex uplink/downlink channel," in *IEEE Information Theory Workshop (ITW)*, Sep. 2013, pp. 1–5.
[9] D. Korpi, Y.-S. Choi, T. Huusari, S. Anttila, L. Talwar, and M. Valkama, "Adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio: Algorithms and RF measurements," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec. 2015.
[10] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.
[11] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.
[12] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.
[13] B. Debaillie, D.-J. van den Broek, C. Lavin, B. van Liempd, E. Klumperink, C. Palacios, J. Craninckx, B. Nauta, and A. Parssinen, "Analog/RF solutions enabling compact full-duplex radios," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1662–1673, Sep. 2014.
[14] M. Isaksson, D. Wisell, and D. Ronnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 1, pp. 348–359, Jan. 2006.
[15] V. Syrjälä, M. Valkama, L. Anttila, T. Riihonen, and D. Korpi, "Analysis of oscillator phase-noise effects on self-interference cancellation in full-duplex OFDM radio transceivers," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 2977–2990, Jun. 2014.

Dani Korpi [S'14] received his B.Sc. and M.Sc. degrees (both with honors) in communications engineering from Tampere University of Technology (TUT), Finland, in 2012 and 2014, respectively. He is currently a researcher in the Department of Electronics and Communications Engineering at TUT, pursuing his D.Sc. (Tech.) degree in communications engineering. His main research interest is the study and development of single-channel full-duplex radios, with a focus on analyzing the RF impairments.

Joose Tamminen received his B.Sc. degree in electrical engineering from Tampere University of Technology (TUT), Finland, in 2016. He is currently continuing his studies towards his M.Sc. degree in electrical engineering at TUT, while working as a research assistant in the Department of Electronics and Communications Engineering. In his work he focuses on design and implementation of embedded systems.

Matias Turunen is currently a research assistant in the Department of Electronics and Communications Engineering at TUT, pursuing his M.Sc degree in electrical engineering. His main research interest is the analog RF cancellation of self-interference in full-duplex radios.

Timo Huusari received his B.Sc. and M.Sc. degrees (both with honors) in electrical engineering from TUT, Finland, in 2012 and 2015, respectively. From 2014 to 2015 he worked for TUT as a research assistant, lead by professor Mikko Valkama, focusing on full-duplex self-interference cancellation devices. In 2015 he joined Intel Labs.

Yang-Seok Choi [S91, M01] received the Ph.D. degree from Polytechnic University, Brooklyn, NY in 2000. Since 1992, he had been with Samsung Electronics, Co., Ltd., in Suwon, Korea, AT&T Labs-Research in NJ, USA and ViVATO, Inc. in Spokane, WA USA. In 2004, he joined Intel Corporation, Hillsboro, OR where he was a Director of Radio Systems Engineering leading Wireless Standards and Technology Development. In 2013, he joined Intel Labs researching on next generation wireless communications.

Lauri Anttila [S06, M11] received the M.Sc. degree and D.Sc. (Tech) degree (with honors) in electrical engineering from Tampere University of Technology (TUT), Tampere, Finland, in 2004 and 2011. Currently, he is a Senior Research Fellow at the Department of Electronics and Communications Engineering at TUT. His research interests are in signal processing for wireless communications, in particular radio implementation challenges in 5G cellular radio and full-duplex radio, flexible duplexing techniques, and transmitter and receiver linearization. He has published over 60 peer reviewed articles in these areas, as well as two book chapters.

Shilpa Talwar is a principal engineer in Intels Wireless Communications Laboratory and leads a research team focused on improving the wireless service experience of 5G subscribers. Prior to Intel, she held several senior technical positions in the wireless industry, working on algorithm and system design for 3G/4G/WiFi networks, satellite communications, and GPS technologies. She graduated from Stanford University with a Ph.D. in applied mathematics and an M.S. in electrical engineering. She is credited with 60+ technical publications and patents.

Mikko Valkama [S'00, M'02, SM'15] received his M.Sc. and Ph.D. degrees (both with honors) from TUT in 2000 and 2001, respectively. In 2003, he worked as a visiting researcher with the Communications Systems and Signal Processing Institute at San Diego State University, California. Currently, he is a full professor and vice-head of the Department of Electronics and Communications Engineering at TUT. His general research interests include communications signal processing, cognitive radio, full-duplex radio, radio localization, and 5G mobile cellular radio systems.

Publication 5

D. Korpi, M. Heino, C. Icheln, K. Haneda, and M. Valkama, "Compact inband full-duplex relays with beyond 100 dB self-interference suppression: Enabling techniques and field measurements," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 2, pp. 960–965, Feb. 2017. DOI: 10.1109/TAP.2016.2632740

Compact Inband Full-Duplex Relays with Beyond 100 dB Self-Interference Suppression: Enabling Techniques and Field Measurements

Dani Korpi, Mikko Heino, Clemens Icheln,
Katsuyuki Haneda, and Mikko Valkama

***Abstract*—In this paper, the self-interference channel and novel enabling techniques for a compact inband full-duplex relay are described and characterized in different operating environments. The full-duplex operation is based on a novel antenna design that uses wavetraps to provide passive isolation of up to 70 dB between the transmit and receive antenna ports. The passive isolation is complemented with novel active RF and digital cancellation stages that further suppress the residual SI to the receiver noise floor. Measurement results of a complete prototype implementation show that the proposed design can achieve an overall SI cancellation performance of over 100 dB even with an ambitious instantaneous bandwidth of 80 MHz. Similar results are obtained both in an anechoic chamber as well as in realistic multipath indoor environments.**

***Index Terms*—Full-duplex, relay, self-interference, isolation, antenna design, digital cancellation, RF measurements**

I. Introduction

Inband full-duplex communications is a recent paradigm shift in the field of wireless communications [1]–[4]. It means that, instead of the traditional time division duplexing (TDD) or frequency division duplexing (FDD), the transmission and reception are done simultaneously on the same center-frequency. The greatest challenge of inband full-duplex communications is the own transmit signal, which is coupling to the receiver, and thereby becomes a significant source of self-interference (SI). Such SI can be several orders of magnitude more powerful than any received signal of interest. Nevertheless, various research groups have already constructed demonstrator implementations that are capable of efficient SI suppression in the receiver, with state-of-the-art results being described in [2], [5]–[7].

One of the most prominent applications for a full-duplex radio is an inband relay [4], [5]. In relaying, the same amount of time is spent for receiving as well as for transmitting, which is well suitable for an inband full-duplex device. Also, a full-duplex relay would transmit and receive continuously on the same frequency band, which means that no changes in the frequency planning or time slot allocations are needed in the network. Thus, inband full-duplex relays could be implemented such that they are entirely transparent from the network perspective.

A trivial way to ensure sufficient isolation between the transmitter and the receiver in a relay is to increase the distance between the transmit and receive antennas [8]. However, this would also increase the size of the relay, which is often not desirable or feasible. In order to construct a compact inband full-duplex relay, more advanced techniques are needed [5], [9]. In this paper, we describe and evaluate a novel architecture based on a compact MIMO-capable antenna design that utilizes wavetraps to provide up to 70 dB of passive isolation between all the transmit and receive antennas. This means that the proposed relay architecture obtains high levels of passive SI suppression while still maintaining a very compact size.

The isolation provided by the antenna design is then complemented by active SI cancellation stages. In this work, two scenarios are considered: (*i*) the received SI is cancelled both in the RF and digital domains, and (*ii*) the received SI is cancelled only in the digital domain. For the digital cancellation stage, a novel nonlinear processing structure combined with adaptive basis function orthogonalization scheme is also presented. The reported measurement results illustrate that both of these active cancellation approaches can suppress the SI to the receiver noise floor, with instantaneous bandwidths of up to 80 MHz. With the used transmit power levels, this translates to a total SI suppression of over 100 dB. When using a narrower bandwidth of 20 MHz, the total amount of SI cancellation in a realistic indoor environment is nearly 110 dB. To the best of our knowledge, these are the highest reported SI cancellation performances for any inband full-duplex relay, measured in true live operating environments.

The proposed antenna design has been previously introduced in [10]. However, there the antenna was characterised only in an anechoic chamber. This paper provides comprehensive SI channel measurements also in realistic multipath environments to validate the antenna design in real use scenarios. Another novelty of this paper is the inclusion of advanced active SI cancellation stages in the relay, which reflect further developments compared to existing prior art presented in [5], [11]. In particular, it is shown that a passive MIMO-capable relay antenna design together with active cancellation only in the digital domain can already bring the SI down to the noise floor. Compared to, e.g., [2], this avoids the need for an analog cancellation circuit board, which can be complex and power consuming, especially in a MIMO system. Furthermore, the proposed relay architecture does not need to resort to spatial SI filtering [4], which means that all the degrees-of-freedom of the MIMO channels can be used for the improvement of desired links.

The rest of this paper is organized as follows. Section II introduces the antenna design and the active RF and digital cancellation solutions. After this, Section III characterizes the antenna performance in different environments through channel measurements. Measured overall SI cancellation performances for an anechoic chamber and a realistic indoor environment are then reported and analysed in Section IV. Finally, Section V concludes the paper.

II. Advanced SI Suppression Solutions

A. High-Isolation Antenna Design

In order to minimize the SI power coupling to the receiver, an antenna with a high TX-RX isolation of 65–70 dB is adopted. The size of the antenna arrangement is similar to a typical WLAN access point and it has two ports on both sides of the structure, making it capable of 2×2 MIMO operation. Inherently, this kind of an antenna structure has a back-to-back TX-RX isolation of 50 dB. However, by adding wavetraps around the edges of the structure, the ground plane currents are suppressed, which improves the isolation by 15–20 dB between all TX and RX port pairs. This is caused by the reduced edge diffraction, which significantly decreases the coupling between the TX and RX sides. The antenna is illustrated in Fig. 1 and detailed in [10].

D. Korpi and M. Valkama are with the Department of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland, e-mail: dani.korpi@tut.fi.

M. Heino, C. Icheln and K. Haneda are with the Department of Radio Science and Engineering, Aalto University School of Electrical Engineering, Espoo, Finland, e-mail: mikko.heino@aalto.fi.

This research work was funded by the Academy of Finland (under the projects #259915, #258364), Tekes (under the TAKE-5 project), Tampere University of Technology Graduate School, Aalto ELEC Doctoral School, Nokia Foundation, HPY Research Foundation, Tuula and Yrjö Neuvo Research Fund, and Emil Aaltonen Foundation.

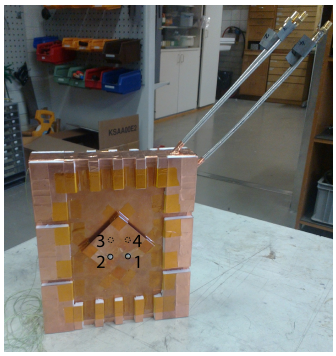


Fig. 1: The measured relay antenna with wavetraps. Ports 3 and 4 are on the other side of the relay.

B. Multitap RF Cancellation

The purpose of the RF canceller is then to provide further SI attenuation on top of the passive antenna isolation before the actual receiver chain. This ensures that the overall received signal does not saturate the receiver low-noise amplifier (LNA) or the analog-to-digital converter (ADC). Essentially, the RF canceller uses the power amplifier (PA) output signal to generate a replica of the observed SI signal in the receiver input, and consequently cancel it out. As described in our earlier work in [11], multiple RF branches or taps can be adopted, all with different RF delays and tunable amplitudes and phases, to facilitate accurate SI regeneration.

The particular RF cancellation solution developed and measured in this work adopts two branches or taps, since this is enough to cancel the direct leakage after the high isolation provided by the antenna. The multipath components of the SI are then dealt with in the digital domain, since cancelling them in the RF domain would require a prohibitively large number of taps in the RF canceller [12]. Hence, when described at baseband equivalent level, the signal after the RF cancellation can be expressed as

$$y_{RF}\left(t\right) = y\left(t\right) - \sum_{n=1}^{2} w_n x_{PA}\left(t-\tau_n\right), \tag{1}$$

where $y\left(t\right)$ is the receiver input signal, $x_{PA}\left(t\right)$ is the PA output signal, w_n is the complex weight of the nth tap, and τ_n is the relative delay of the input signal of the nth tap. The correct amplitude and phase settings for the two taps are obtained by a self-adaptive control algorithm, implementing an LMS type learning rule. This can be described as

$$w_n \leftarrow w_n + \mu_{RF} \int \left(x_{PA}^{IQ}\left(t-\tau_n\right)\right)^{*} y_{RF}^{IQ}\left(t\right) dt, \tag{2}$$

where μ_{RF} is the RF canceller step size, $x_{PA}^{IQ}\left(t\right)$ is the IQ demodulated baseband observation of the PA output signal, and $y_{RF}^{IQ}\left(t\right)$ is the IQ demodulated RF canceller output signal. This type of a closed-loop control system allows the RF canceller to learn and track the coupling channel in real time and thereby ensure that the power level of the SI entering the actual receiver chain is not too high. Compared to our earlier work and results described in [5], [11], where purely analog control loop was adopted, the RF canceller developed in this work builds on a digital control loop, meaning that the learning of the weights, described by (2), is done fully in the digital domain. This obviously allows for a more flexible implementation of the RF canceller.

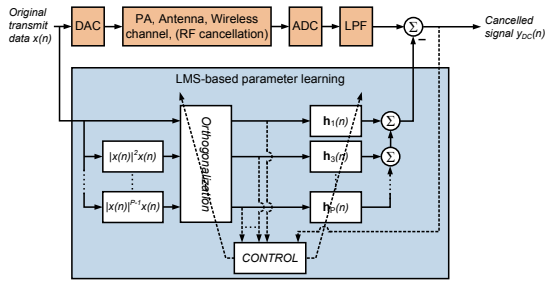


Fig. 2: A general illustration of the LMS based adaptive nonlinear digital canceller, incorporating a parallel Hammerstein model for a wideband nonlinear power amplifier. Here, $\mathbf{h}_p(n)$ denotes the coefficients of the orthogonalized pth-order basis function, which together comprise the combined coefficient vector $\mathbf{h}(n)$.

C. Nonlinear Digital Cancellation

To finally suppress the residual SI to the level of the receiver noise floor, further cancellation in the digital domain is still typically required [3], [5]. Digital cancellation can be performed using the original transmit data as the reference signal, which is then properly filtered with an estimate of the overall SI channel. In this case, the overall SI channel includes the effects of the transmitter, the wireless channel, the RF canceller, and the receiver. In terms of the performance of the digital canceller, obtaining an accurate estimate of the SI channel is the most crucial aspect.

A particularly challenging issue in digital SI cancellation is the nonlinear nature of the transmitter PA [2], [3]. Namely, when utilizing a power-efficient PA, the relationship between the observed SI and the original transmit signal is no more linear. This means that a nonlinear signal model must be used in the digital canceller, as shown already in [2], [3].

In this work, a nonlinear digital canceller combined with LMS-based parameter learning is developed. Assuming that the transmitter PA is the only significant source of nonlinear distortion, a parallel Hammerstein (PH) model can be used to accurately model the observed SI [13]. As a starting point, let us define the so-called basis function vector as

$$\mathbf{\Phi}(n) = \begin{bmatrix}\phi_1(x(n)) & \phi_3(x(n)) & \cdots & \phi_P(x(n))\end{bmatrix}^T, \tag{3}$$

where $\phi_p\left(x(n)\right) = |x(n)|^{p-1}\, x(n)$ is the pth-order nonlinear basis function, $x(n)$ is the original baseband transmit signal, and P is the nonlinearity order of the signal model. The basic idea behind the PH model is to allow for estimating the effective SI channel separately for each basis function, incorporating also memory. However, because the different nonlinear terms are typically highly correlated, resulting in poor convergence of LMS-type parameter learning algorithms, they must be efficiently orthogonalized. This can be done by using a novel adaptive orthogonalization procedure alongside the actual LMS parameter learning, as described below. The overall adaptive nonlinear digital cancellation stage is illustrated in Fig. 2.

Denoting first the instantaneous orthogonalization matrix at time index n by $\mathbf{W}\left(n\right)$, the basis function orthogonalization step can be expressed as

$$\widetilde{\mathbf{\Phi}}(n) = \mathbf{W}\left(n\right)\mathbf{\Phi}(n). \tag{4}$$

The orthogonalization matrix is then updated during each iteration

using the following rule [14, Chapter 4]:

$$\mathbf{W}(n+1) = \mathbf{W}(n) - \mu_{ort}\left(\widetilde{\mathbf{\Phi}}(n)\widetilde{\mathbf{\Phi}}^{H}(n) - \mathbf{I}\right)\mathbf{W}(n), \quad (5)$$

where μ_{ort} is the learning parameter for the adaptive orthogonalizer, $(\cdot)^H$ denotes the Hermitian transpose, and $\mathbf{I}$ is the identity matrix. The orthogonalization matrix is typically initialized as $\mathbf{W}(0) = \mathbf{I}$.

Then, the actual LMS-based learning step is carried out for the effective SI channel parameters, corresponding to the orthogonalized nonlinear basis function samples. Utilizing (4), the input of the combined LMS filter $\mathbf{h}(n)$ can be written as

$$\mathbf{u}(n) = \left[\widetilde{\mathbf{\Phi}}(n+M_1)^T \quad \widetilde{\mathbf{\Phi}}(n+M_1-1)^T \quad \cdots \quad \widetilde{\mathbf{\Phi}}(n-M_2)^T\right]^T \quad (6)$$

where M_1 and M_2 are the amounts of pre-cursor and post-cursor memory, respectively. Then, the canceller output signal at time-instant n can be written as

$$y_{DC}(n) = y_{RF}(n) - \mathbf{h}(n)^H\mathbf{u}(n), \quad (7)$$

where $y_{RF}(n)$ denotes the received (and possibly RF cancelled) signal in the digital domain. The final step in the LMS parameter learning procedure is then to update the coefficients $\mathbf{h}(n)$ based on the cancelled sample. The update rule can be expressed as

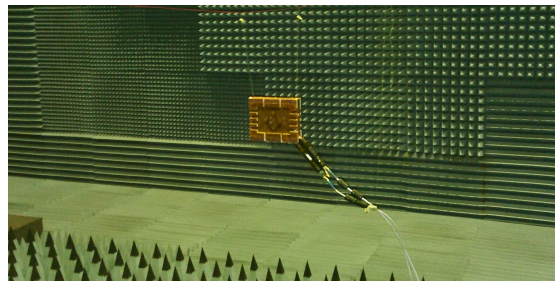
$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mathbf{\Lambda}y^*_{DC}(n)\mathbf{u}(n), \quad (8)$$

where $\mathbf{\Lambda}$ is a diagonal matrix consisting of the step sizes for each nonlinear basis function and $(\cdot)^*$ denotes the complex conjugate. The coefficient vector is commonly initialized as $\mathbf{h}(0) = \mathbf{0}$ if no further side-information is available. This type of an LMS-based adaptive nonlinear digital canceller allows for accurate regeneration of the residual SI under a nonlinear PA, as well as efficient parameter tracking under time-varying SI channel conditions.

III. Antenna Measurements

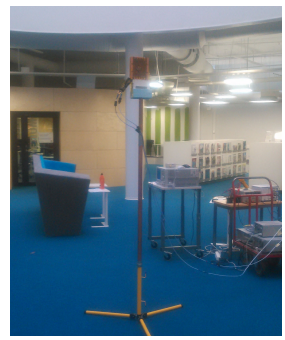
Previously in [10], the antenna performance was only measured in an anechoic chamber, i.e., under ideal conditions without reflections from the environment that can easily couple the TX signal to the RX antenna. In this section, the performance of the antenna is studied in different real multipath environments. Especially, if there are reflecting surfaces near the antenna, the reflections from the environment can in general decrease the overall isolation [7], [8]. An interesting aspect is the ratio between the directly coupled SI power and the reflected SI power, which is analogous to the Rician K-factor of fading channels, since the antennas with very high inherent TX-RX isolation, such as the prototype antenna used in this paper, only attenuate the direct SI coupling. It is expected that at some point, when the environment becomes sufficiently reverberant, the reflected SI power starts to dominate, thereby introducing an environment-dependent upper limit on the achievable antenna isolation.

The coupling for co-polarized TX-RX port pair 2-4 of the antenna, measured in the different environments shown in Fig. 3, is presented in Fig. 4. It can clearly be observed that the multipath propagation decreases the attainable isolation, while causing ripples in the S-parameters due to destructive and constructive interference of the different multipath components. For reference, Fig. 4 also shows the return loss of the antenna, from where it can be observed that its bandwidth is roughly 150 MHz for a maximum return loss of -10 dB.



(a)

(b)



(c)

(d)

Fig. 3: The measurement setup (a) in the chamber, (b) outdoors, (c) indoors I, and (d) indoors II.

In order to investigate more closely the isolation of the antenna in the different environments, the second column of Table I shows the mean isolation $(\frac{1}{N}\sum_{f=f_{\min}}^{f_{\max}}|S_{42}|^2)^{-1}$ for the frequencies between 2.55–2.65 GHz. From there, it can be seen that the measured isolation in the anechoic chamber and outdoors stays roughly the same. Thereby, in the outdoors case, the environment does not affect the obtained isolation much, owing to the sparsity of scatterers around the antenna. However, in the first indoor case, the isolation over a 100 MHz bandwidth is degraded by 4.4 dB when compared to the anechoic chamber. In the second indoor case, the isolation degradation is 7.1 dB. This degradation can be explained by the increase in the multipath SI power in the more reverberant environments.

Figure 5 shows the impulse responses calculated with non-windowed IFFT from the S-parameters for the frequency range of 2.5–2.7 GHz with delay resolution of 5 ns. The increase in the power of late reflections in the more reverberant indoor environments is evident. Moreover, since the reflecting surfaces

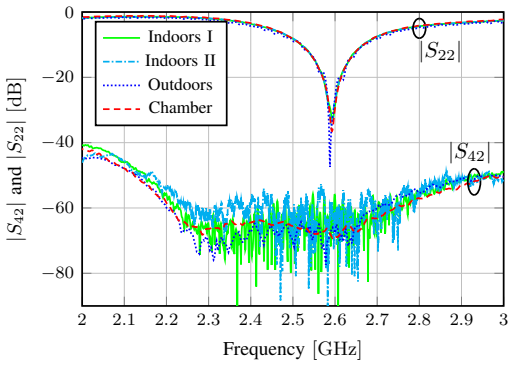


Fig. 4: The S-parameters of the antenna in the different environments.

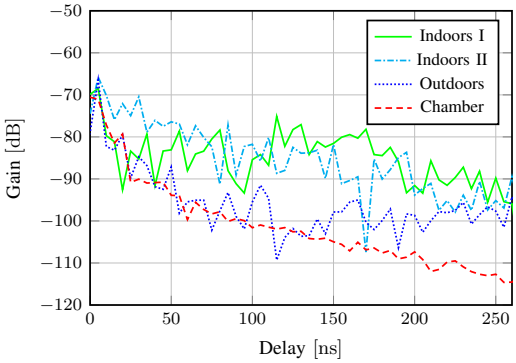


Fig. 5: SI channel impulse responses in the different environments.

are closer to the antenna in the second indoor environment, the early reflected SI power is higher than in the first indoor case. The direct coupling in the antenna structure remains roughly the same in each case.

From Fig. 5, the power differences between the direct coupling and late reflections can be calculated for each environment. The rightmost column of Table I compares the power associated with the first two delay taps (0–7.5 ns), containing the direct coupling and the short delay by the cables in Fig. 1, to the power received after 7.5 ns. In the outdoor and chamber results, it can be seen that the received direct SI dominates. However, in the indoor environments, the power received from the multipath reflections is actually higher than the direct coupling. This indicates that, in closed environments, the ability of the antenna to reduce the SI is limited. In other words, this means that the current relay design is already good enough for such environments, since increasing the antenna isolation further would not help to reduce the SI.

IV. Overall Cancellation Results

In order to evaluate the overall performance of the proposed full-duplex relay architecture, RF field measurements are next performed on a complete demonstrator setup. The whole system is evaluated both in the anechoic chamber as well as in the first indoor deployment scenario, where the antenna is located in a university building. In the chamber measurements, also an RF canceller is employed, while in the indoor measurements it is left out. This is done in order to put maximum processing burden on the developed nonlinear digital SI canceller and thus test its full cancellation power.

TABLE I: The mean antenna isolation from 2.55 to 2.65 GHz, together with the ratio of direct (≤ 7.5 ns) and multipath (> 7.5 ns) SI.

Environment	**Mean isolation**	$\frac{P_{\mathrm{SI}}(\mathrm{delay} \leq 7.5\,\mathrm{ns})}{P_{\mathrm{SI}}(\mathrm{delay} > 7.5\,\mathrm{ns})}$
Chamber	67.2 dB	5.7 dB
Outdoors	68.1 dB	8.1 dB
Indoors I	62.8 dB	−0.3 dB
Indoors II	60.1 dB	−2.5 dB

A. Chamber Measurements

Let us first investigate the overall performance of the full-duplex relay under the ideal conditions of the anechoic chamber. The measurement setup is as shown in Fig. 3(a), where the antenna itself is suspended in the air to avoid any undesired effects caused by supporting structures. The essential measurement parameters are listed in Table II.

In the measurements, a National Instruments PXIe-5645R vector signal transceiver (VST) is used both as the transmitter and the receiver. The transmit signal is an LTE waveform of a specified bandwidth and an output power of -5 dBm, centred at 2.56 GHz. The VST output signal is then further amplified by an external PA with a gain of 36 dB, after which the transmit signal is divided between the RF canceller and the antenna. Accounting for all the losses in the transmission path, the final transmit power is in the order of 29 dBm for all the considered bandwidths.

The signal that is received by the antenna is first routed to the RF canceller board, during which it is attenuated by 5 dB due to the cable and insertion losses. The RF canceller contains two branches with excess delays of 5 ns and 7.5 ns, respectively, targeting to suppress primarily the direct SI component. Note that, as already stated earlier, in order to suppress the various multipath components evident in Fig. 5 in the RF domain, an impractical number of taps would be required, and hence they can be attenuated more efficiently in the digital domain. The delays of the two taps are chosen based on the impulse response of the direct coupling through the antenna, such that as much of the energy of the SI as possible is between the taps while their mutual delay is minimized. For more information regarding the interplay between the tap delays and RF cancellation performance, refer to [12]. After RF cancellation, the signal is fed to the VST-receiver, which records the I- and Q-samples for digital post processing.

The digital cancellation procedure is then performed in the host processor using the recorded samples. The digital canceller utilizes adaptive nonlinear modeling of the residual SI waveform and the adaptive orthogonalization procedure, using the parameters specified in Table II. Both the nonlinearity order and the number of pre- and post-cursor taps are chosen based on experimental evaluation such that the cancellation performance is maximized. The required nonlinearity order stems from the characteristics of the used low-cost PA, while the required number of taps is mostly determined by the nature of the wireless coupling channel as well as by the various memory effects introduced by the transceiver itself. Learning parameters in the order of $10^{-3} \ldots 10^{-6}$ are then used for the adaptation of the orthogonalization matrix and the coefficients of the nonlinear basis functions. In addition, the cancellation performance of a regular linear digital canceller is also

TABLE II: The relevant parameters used in the overall cancellation measurements.

Parameter	Chamber	Indoors
Bandwidth	20/40/80 MHz	20/40/80 MHz
Transmit power	29 dBm	24 dBm
Center frequency	2.56 GHz	2.56 GHz
PA gain	36 dB	36 dB
RF canceller	Yes	No
Highest nonlinearity order (P)	11	11
Number of pre-cursor taps M_1	50	50
Number of post-cursor taps M_2	50	50

evaluated and reported for reference. The linear digital canceller is obtained as a special case of the described nonlinear digital canceller simply by setting the nonlinearity order $P = 1$.

The signal spectra after all the different cancellation stages are shown in Fig. 6 for bandwidths of 20, 40, and 80 MHz. It can be observed that the isolation provided by the considered relay antenna is in the order of 70 dB for all the bandwidths, even after taking into account the losses in the reception path. This result is well in line with the observations made in Section III. After this, the SI is further suppressed by the two-tap RF canceller, which can attenuate the SI by 7 to 17 dB, depending on the bandwidth. The relatively low amount of RF cancellation is explained by the extraordinarily high isolation provided already by the antenna, which means that the direct SI leakage is rather weak even before the RF canceller.

The final cancellation stage is then the digital canceller, which in this case is in fact able to attenuate the residual SI nearly perfectly, even with the 80 MHz transmit signal, as can be observed in Fig. 6(c). The slight difference between the signal power after the digital cancellation and the theoretical receiver noise floor can be partially attributed to the RF canceller circuitry and the small amount of distortion it produces. In particular, with the wider bandwidths, the RF canceller is producing some spurious tones, as can be observed in Figs. 6(b) and 6(c), and this somewhat contributes to the power of the residual SI. Nevertheless, even after accounting for all the losses in the transmitter and the receiver, the overall amount of SI cancellation is still in the order of 110 dB for 20 MHz bandwidth, and roughly 100 dB for 80 MHz bandwidth. These are one of the highest reported SI cancellation performances for such wide bandwidths [2], [5]. Furthermore, Fig. 6 shows that reaching these cancellation levels requires nonlinear modeling in the digital domain, since the linear digital canceller is not capable of sufficiently accurate SI regeneration.

B. Indoor Measurements

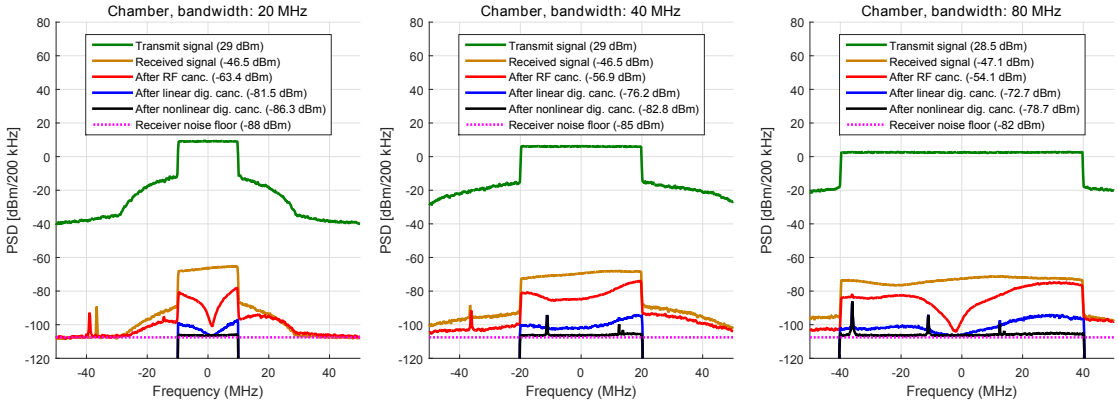
The overall indoor performance of the full-duplex relay is evaluated in the first indoor environment, the measurement location being shown in Fig. 3(c). A similar measurement setup was used as in the chamber measurements, with all the essential parameters listed again in Table II. The only difference is that now the transmission is done with a lower VST output power of -10 dBm, which results in an overall transmit power of roughly 24 dBm. In addition, due to the high amount of antenna isolation and the presence of various multipath components with large delays, the RF canceller is not used in this case, since it is not capable of providing further attenuation under such very long echoes. For this reason, and to test the performance boundaries of digital SI cancellation, the forthcoming results incorporate only the antenna isolation and the nonlinear digital canceller.

The signal spectra after the different cancellation stages for the three different transmission bandwidths are shown in Fig. 7. It can be observed that, even under the realistic deployment scenario, the antenna provides an isolation in the order of 65 dB for bandwidths up to 80 MHz. The nonlinear digital canceller can then attenuate the SI further by around 40 dB, pushing the SI practically to the receiver noise floor with all bandwidths. Due to the relatively high transmit power, the PA produces again a significant amount of nonlinear distortion, which means that the linear digital canceller is not capable of generating a sufficiently accurate cancellation signal. In Fig. 7, this is evident from the lower amount of cancellation obtained with the linear canceller. Taking into account the overall receiver loss of 4 dB, the total amount of cancellation is thereby 106 dB, 103 dB, or 100 dB for bandwidths of 20 MHz, 40 MHz, or 80 MHz, respectively. Note that the level of the residual SI in the 40 and 80 MHz cases is somewhat increased by the spurious tone located around 13 MHz, and hence the true amount of SI cancellation can be assumed to be slightly more. This spurious tone is likely caused by the PXIe-5645R VST itself, made visible by the high transmit power and the consecutive large amount of SI cancellation.

When comparing the measurement results from the chamber and from indoors, it can be seen that the overall cancellation performance is largely similar. This holds despite the effective coupling channel being highly frequency selective in the indoor measurements, as can be observed in Fig. 7. This frequency selectivity results in the RF canceller not performing optimally, since it has been designed for scenarios with a dominant signal component present. Hence, even though it provides a reasonable amount of cancellation in the chamber measurements, it cannot cancel the SI in the indoor measurements. Even with this challenge, however, the nonlinear digital canceller can provide the necessary amount of SI suppression also in the indoor measurements, and the residual SI is consequently attenuated to the receiver noise floor. Hence, even though the analog SI cancellation is absent in the indoor measurements, the overall cancellation performance is still nearly the same, indicating the high robustness and excellent cancellation power of the digital canceller. All in all, these integrated measurement results show that, in terms of SI cancellation performance, the implemented relay architecture is already capable of achieving true full-duplex operation even for very wideband systems in challenging true environments.

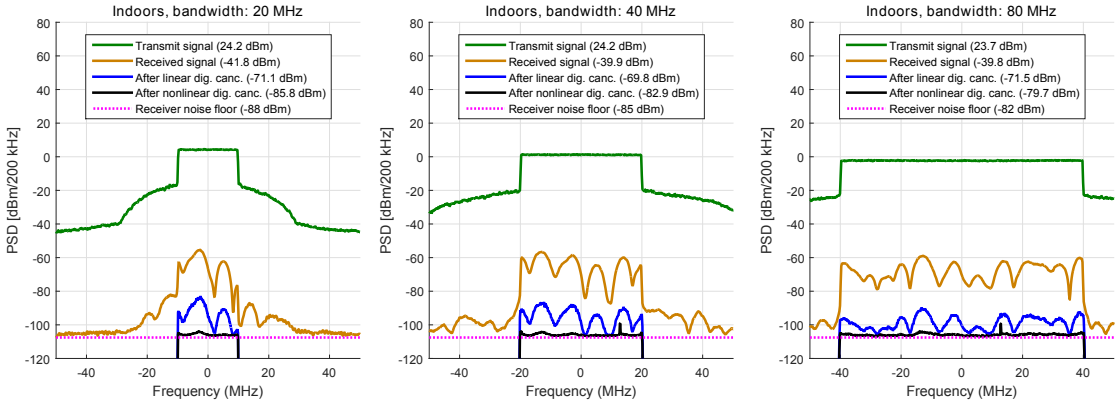
V. Conclusions

In this paper, key enabling techniques for a compact inband full-duplex relay were described and evaluated. First, the operation of a high-isolation relay antenna was characterized in four different environments. It was concluded that there is a limit on how much isolation can be obtained with high-isolation relay antennas in reverberant environments. In order to fully suppress the residual self-interference, further active cancellation stages in the analog and/or digital domain were then introduced. Together with these additional cancellers, it was shown that the described full-duplex relay architecture can achieve overall self-interference cancellation performances of over 100 dB in realistic environments, even with only digital cancellation and ambitious bandwidths up to 80 MHz. This means that the self-interference is canceled to the receiver noise floor even with transmit powers in the order of 30 dBm, despite a heavily nonlinear transmitter power amplifier.



(a) (b) (c)

Fig. 6: The signal spectra after the different cancellation stages when measured in an anechoic chamber using (a) 20 MHz, (b) 40 MHz, and (c) 80 MHz LTE waveforms.



(a) (b) (c)

Fig. 7: The signal spectra after the different cancellation stages when measured indoors using (a) 20 MHz, (b) 40 MHz, and (c) 80 MHz LTE waveforms.

REFERENCES

[1] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 10, pp. 1637–1652, Oct. 2014.

[2] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. SIGCOMM'13*, Aug. 2013, pp. 375–386.

[3] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.

[4] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[5] M. Heino et al., "Recent advances in antenna design and interference cancellation algorithms for in-band full-duplex relays," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 91–101, May 2015.

[6] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[7] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[8] K. Haneda, E. Kahra, S. Wyne, C. Icheln, and P. Vainikainen, "Measurement of loop-back interference channels for outdoor-to-indoor full-duplex radio relays," in *Proc. IEEE 4th European Conference on Antennas and Propagation*, 2010, pp. 1–5.

[9] E. Foroozanfard, O. Franek, A. Tatomirescu, E. Tsakalaki, E. De Carvalho, and G. F. Pedersen, "Full-duplex MIMO system based on antenna cancellation technique," *IET Electron. Lett.*, vol. 50, no. 16, pp. 1116–1117, Jul. 2014.

[10] M. Heino, S. Venkatasubramanian, C. Icheln, and K. Haneda, "Design of wavetraps for isolation improvement in compact in-band full-duplex relay antennas," *IEEE Trans. Antennas Propag.*, vol. 64, no. 3, pp. 1061–1070, Mar. 2016.

[11] T. Huusari, Y.-S. Choi, P. Liikkanen, D. Korpi, S. Talwar, and M. Valkama, "Wideband self-adaptive RF cancellation circuit for full-duplex radio: Operating principle and measurements," in *Proc. IEEE 81st Vehicular Technology Conference*, May 2015, pp. 1–7.

[12] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.

[13] M. Isaksson, D. Wisell, and D. Ronnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 1, pp. 348–359, Jan. 2006.

[14] A. Cichocki and S.-i. Amar, *Adaptive Blind Signal and Image Processing*. John Wiley & Sons, 2002.

Publication 6

D. Korpi, L. Anttila, and M. Valkama, "Nonlinear self-interference cancellation in MIMO full-duplex transceivers under crosstalk," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, Feb. 2017. DOI: 10.1186/s13638-017-0808-4

Korpi *et al. EURASIP Journal on Wireless Communications and Networking* (2017) 2017:24
DOI 10.1186/s13638-017-0808-4

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Nonlinear self-interference cancellation in MIMO full-duplex transceivers under crosstalk

Dani Korpi*, Lauri Anttila and Mikko Valkama

Abstract

This paper presents a novel digital self-interference canceller for an inband multiple-input-multiple-output (MIMO) full-duplex radio. The signal model utilized by the canceller is capable of modeling the in-phase quadrature (IQ) imbalance, the nonlinearity of the transmitter power amplifier, and the crosstalk between the transmitters, thereby being the most comprehensive signal model presented thus far within the full-duplex literature. Furthermore, it is also shown to be valid for various different radio frequency (RF) cancellation solutions. In addition to this, a novel complexity reduction scheme for the digital canceller is also presented. It is based on the widely known principal component analysis, which is used to generate a transformation matrix for controlling the number of parameters in the canceller. Extensive waveform simulations are then carried out, and the obtained results confirm the high performance of the proposed digital canceller under various circuit imperfections. The complexity reduction scheme is also shown to be capable of removing up to 65% of the parameters in the digital canceller, thereby significantly reducing its computational requirements.

Keywords: Full-duplex, MIMO, Self-interference, RF impairments, Crosstalk

1 Introduction

Inband full-duplex communications is a promising candidate technology for further improving the spectral efficiency of the next generation wireless systems, such as the 5G networks [1–11]. The basic idea behind it is to transmit and receive at the same time at the same center-frequency, thereby in principle doubling the spectral efficiency. The drawback of such inband full-duplex operation is the own transmit signal, which is coupling to the receiver and becomes an extremely powerful source of self-interference (SI). The most significant challenge in implementing inband full-duplex radios in practice is thereby the development of SI cancellation solutions, which are capable of removing the SI in the receiver. There are already reports of various demonstrator implementations, which achieve relatively high SI cancellation performance, thereby allowing for true inband full-duplex operation [1–3, 6, 7, 11–14].

Moreover, in order to meet the high throughput requirements of the future wireless networks, it is inevitable that the inband full-duplex concept must be combined with MIMO capabilities in the transceivers [7, 12–19]. This obviously results in a higher physical layer capacity, but it also requires more elaborate SI cancellation solutions. In particular, in a MIMO transceiver, the observed SI signal in each receiver consists of a combination of all the transmit signals, which means that also the SI cancellers must have all of the transmit signals available. Furthermore, in order to perform SI cancellation, the coupling channels between all the transmitters and receivers must be estimated, which results in a somewhat more demanding SI cancellation procedure. Nevertheless, this increased complexity is justified by the higher physical layer throughputs.

Especially the complexity of the RF canceller is heavily affected by the number of transmitters and receivers [7, 15]. For an $N_T \times N_R$ MIMO transceiver, the RF canceller requires at least $N_T N_R$ cancellation paths, or even more if using a multi-tap solution [7, 20]. This number can be somewhat decreased by using auxiliary transmitters to upconvert digitally generated cancellation signals, since then only N_R cancellation paths are required. However, the drawback of this solution is obviously the need for additional RF transmitters, as well as the fact that the

*Correspondence: dani.korpi@tut.fi
Department of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland

digitally generated cancellation signals do not include any of the transmitter-induced impairments, which thereby remain unaffected by this type of an RF cancellation solution [1]. Another possible solution for decreasing the complexity of RF cancellation in the context of very large transmit antenna arrays is to use beamforming to form nulls in the receive antennas [4, 21], which might even allow for completely omitting RF cancellation. In typical MIMO devices, however, the increase in the RF cancellation complexity is more or less inevitable.

Also, the complexity of digital SI cancellation is somewhat increased under MIMO operation, but it is obviously more straight-forward to process several SI signals in the digital domain. In particular, more computational resources are needed to estimate all the channel responses between the several transmitters and receivers, but no additional RF hardware is required. However, having several transmit chains on a single chip introduces another issue from the perspective of the digital canceller: the crosstalk between the transmitters, which occurs both before and after the power amplifiers (PAs) [22–28]. This phenomenon is illustrated in Fig. 1 for an example case of three transmitters. What makes this an especially cumbersome issue is the fact that typically the PAs introduce significant nonlinear distortion [3, 29]. This, on the other hand, means that nonlinear modeling of the SI is required in the digital canceller, which is very challenging if the PA input is in fact a linear combination of all the original transmit signals, as is the case under crosstalk [26]. Nevertheless, it is still necessary to model the crosstalk, since otherwise the accuracy of the regenerated SI signal is not sufficiently high. This is especially crucial for the emerging massive MIMO transceivers, where the large amount of transmit chains calls for a high level of integration, which results in more leakage between the transmission paths [28]. Hence, the increase in computational complexity caused by the crosstalk modeling must be tolerated in order to obtain sufficient levels of SI cancellation also under MIMO operation.

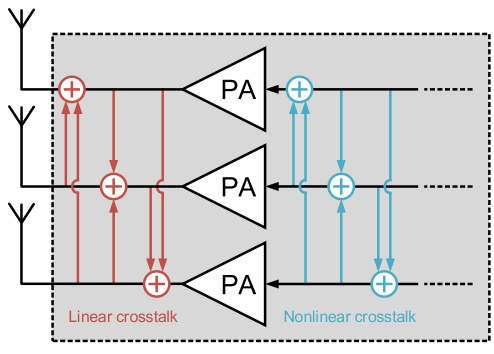


Fig. 1 An illustration of the crosstalk phenomenon in a three-antenna MIMO transmitter, where crosstalk occurs both before and after the PAs. The former is typically referred to as *nonlinear crosstalk*, while the latter is called *linear crosstalk*

In this article, we present a general signal model for the observed SI in the digital domain under a scenario where there is crosstalk between the transmit chains before and after nonlinear PAs. Moreover, it is shown that the signal model can be applied to various different RF cancellation solutions. The presented comprehensive signal model, which shows the effect of the crosstalk in terms of the original transmit signals, is then used as a basis for a high-performance digital SI canceller. The IQ imbalance occurring both in the transmitters and in the receivers is also included in the signal model, since it is typically one of the dominant sources of distortion in a practical transceiver, alongside with the PA-induced nonlinearities [30].

Furthermore, to address the increase in the computational complexity due to the MIMO operation and crosstalk modeling, a novel principal component analysis (PCA)-based solution is proposed, which can be used to control the complexity of the signal model. In particular, PCA processing is used to identify the insignificant terms in the observed SI signal, which are then omitted in the further cancellation processing. This results in a significant reduction of the unknown parameters that must be estimated, which obviously decreases the computational requirements of the digital SI canceller. Moreover, since the most dominant SI terms are retained by such processing, there is no essential degradation in the cancellation performance. To the best of our knowledge, such complexity reduction schemes have not been previously proposed in the context of SI cancellation solutions.

The detailed list of novel contributions in this paper is as follows:

- We derive the most comprehensive MIMO signal model for the observed SI presented so far in the literature. It covers various RF cancellation scenarios, while also modeling the crosstalk between the transmitters under low-cost nonlinear PAs and IQ imbalance.
- We propose a novel nonlinear digital SI canceller, which utilizes the aforementioned advanced signal model.
- We propose a novel complexity reduction scheme based on PCA, which can be used to control the computational complexity of the digital canceller, while minimizing the decrease in the cancellation performance.

- We present numerical results, which illustrate various aspects of the proposed digital SI cancellation solution with realistic waveform simulations.

The rest of this article is organized as follows. In Section 2, the MIMO signal model is derived. Then, in Section 3, the actual nonlinear digital SI canceller is presented, alongside with the parameter estimation procedure and the PCA-based complexity reduction scheme. After this, in Section 4, the proposed digital SI cancellation solution is evaluated with realistic waveform simulations. Finally, the conclusions are drawn in Section 5.

2 Baseband equivalent signal modeling

In this section, we build a complete SI channel model for a MIMO full-duplex device, including the effects of transmitter impairments (PA nonlinearity, IQ imbalance, and transmitter crosstalk), the linear MIMO SI channel, and RF cancellation. In the forthcoming analysis, the nonlinearities produced by the digital-to-analog and analog-to-digital converters (DACs and ADCs) [31], alongside with phase noise, are omitted from the signal model for simplicity, although phase noise is still included in the reported simulation results.

An illustration of the considered full-duplex MIMO transceiver is given in Fig. 2, with two alternative RF cancellation solutions. In particular, the RF cancellation can be done either by utilizing the PA output signals, or by generating the cancellation signals in the digital domain and upconverting them with the help of auxiliary transmitters. In the forthcoming analysis, both of these options are considered. Furthermore, in Fig. 2, the transceiver is shown to have separate transmit and receive antennas only for illustrative purposes, since the same signal model can also be applied to a case where each antenna is shared between a transmitter and a receiver [32]. Hence, the forthcoming analysis is directly applicable also to a shared-antenna architecture. Note that, for notational simplicity, the actual received signals of interest and additive noise are not included in the following presentation.

2.1 Power amplifier and IQ modulator models with crosstalk

Let us denote the baseband signal of transmitter j ($j = 1, 2, \ldots, N_T$) by $x_j(n)$. The output signal of a frequency-independent IQ modulator model is [33]

$$x_j^{IQM}(n) = K_{1,j}x_j(n) + K_{2,j}x_j^*(n) \tag{1}$$

with $K_{1,j} = 1/2(1 + g_j \exp(j\varphi_j)), K_{2,j} = 1/2(1 - g_j \exp(j\varphi_j))$, where g_j, φ_j are the gain and phase imbalance parameters of transmitter j. Notice that under typical circumstances $|K_{1,j}| \gg |K_{2,j}|$. The magnitude of the IQ image component, represented by the conjugated signal term in (1), can be characterized with the image rejection ratio (IRR) as $10\log_{10}\left(|K_{1,j}|^2/|K_{2,j}|^2\right)$.

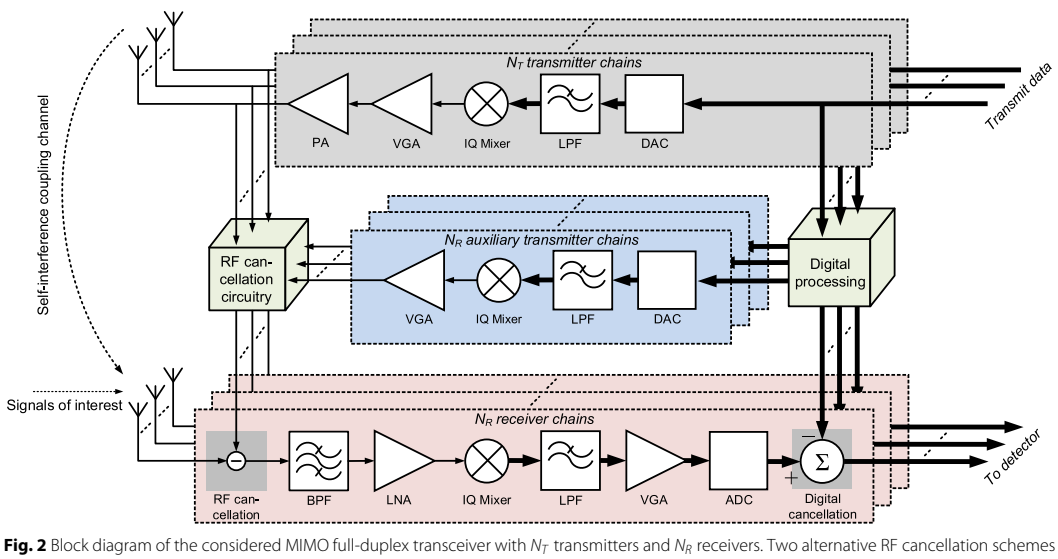


Fig. 2 Block diagram of the considered MIMO full-duplex transceiver with N_T transmitters and N_R receivers. Two alternative RF cancellation schemes are illustrated in the diagram: one utilizing directly the transmitter output signals and one utilizing auxiliary transmitters with digitally generated RF cancellation signals

The response of the PA is approximated using the widely known parallel Hammerstein (PH) model, given for transmitter j as [34]

$$x_j^{PA}(n) = \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{m=0}^{M} h_{p,j}(m)\psi_p(x_{j,in}(n-m)), \tag{2}$$

where $x_{j,in}(n)$ is the PA input signal, the basis functions are defined as

$$\psi_p(x(n)) = |x(n)|^{p-1}x(n) = x(n)^{\frac{p+1}{2}}x^*(n)^{\frac{p-1}{2}} \tag{3}$$

and $h_{p,j}(n)$ denote the impulse responses of the PH branches for transmitter j, while M and P denote the memory depth and nonlinearity order of the PH model, respectively [34–36]. The PH nonlinearity is a widely used nonlinear model for direct as well as inverse modeling of PAs [34–37].

Due to the crosstalk occurring before each PA, referred to as *nonlinear crosstalk*, the input signal $x_{j,in}(n)$ can be written as

$$x_{j,in}(n) = \sum_{i=1}^{N_T} \alpha_{ij} x_i^{IQM}(n), \tag{4}$$

where α_{ij} is the crosstalk coefficient between the ith and jth transmitter chains, and $\alpha_{jj} = 1\,\forall\, j$. In other words, as a result of the crosstalk occurring before the PAs, each PA input signal is in fact a linear combination of all the different transmit signals. The crosstalk phenomenon is illustrated for an example case of three transmitters in Fig. 1, where both the nonlinear and linear crosstalk are shown. Inserting now (1) into (4), we can rewrite the PA input signal as

$$\begin{aligned} x_{j,in}(n) &= \sum_{i=1}^{N_T} \alpha_{ij}\left(K_{1,i}x_i(n) + K_{2,i}x_i^*(n)\right) \\ &= \sum_{i=1}^{N_T} \alpha_{1,ij}x_i(n) + \sum_{i=1}^{N_T} \alpha_{2,ij}x_i^*(n), \end{aligned} \tag{5}$$

where $\alpha_{1,ij} = \alpha_{ij}K_{1,i}$ and $\alpha_{2,ij} = \alpha_{ij}K_{2,i}$.

Using (5), the signal at the PA output can be written as follows:

$$\begin{aligned} x_j^{PA}(n) &= \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{m=0}^{M} h_{p,j}(m)x_{j,in}(n-m)^{\frac{p+1}{2}}x_{j,in}^*(n-m)^{\frac{p-1}{2}} \\ &= \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{m=0}^{M} h_{p,j}(m) \sum_{k_0=0}^{\frac{p+1}{2}} \binom{\frac{p+1}{2}}{k_0} \\ &\quad \times \left(\sum_{i=1}^{N_T} \alpha_{1,ij}x_i(n-m)\right)^{\frac{p+1}{2}-k_0} \left(\sum_{i=1}^{N_T} \alpha_{2,ij}x_i^*(n-m)\right)^{k_0} \\ &\quad \times \sum_{l_0=0}^{\frac{p-1}{2}} \binom{\frac{p-1}{2}}{l_0} \left(\sum_{i=1}^{N_T} \alpha_{2,ij}^*x_i(n-m)\right)^{\frac{p-1}{2}-l_0} \\ &\quad \times \left(\sum_{i=1}^{N_T} \alpha_{1,ij}^*x_i^*(n-m)\right)^{l_0} \end{aligned} \tag{6}$$

It can be further modified by expanding all the integer powers of the sum signals as shown in the Appendix, which gives

$$\begin{aligned} x_j^{PA}(n) = \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{q_0=0}^{p} \sum_{q_1=0}^{p-q_0} \cdots \sum_{q_{N_T-1}=0}^{p-q_0-\cdots-q_{N_T-2}} \sum_{r_1=0}^{q_0} \sum_{r_2=0}^{q_0-r_1} \cdots \\ \sum_{r_{N_T-1}=0}^{q_0-r_1-\cdots-r_{N_T-2}} \sum_{m=0}^{M} h_{p,j,q_0,\ldots,r_{N_T-1}}(m) \\ \times x_1(n-m)^{q_1}x_2(n-m)^{q_2}\cdots \\ x_{N_T}(n-m)^{p-\sum_{i=0}^{N_T-1}q_i} \\ \times x_1^*(n-m)^{r_1}x_2^*(n-m)^{r_2}\cdots \\ x_{N_T}^*(n-m)^{q_0-\sum_{i=1}^{N_T-1}r_i} \end{aligned} \tag{7}$$

where $h_{p,j,q_0,\ldots,r_{N_T-1}}(m)$ are the coefficients for the basis function of the form $\prod_{i=1}^{N_T} x_i(n)^{a_i}x_i^*(n)^{b_i}$ such that $\sum_{i=1}^{N_T}(a_i+b_i) = p$. This signal model is of similar form as the one presented in [26], with the exception that the model in (7) also incorporates the effect of IQ imbalance and is thus more complete.

In order to simplify (7), it can be noted that, for the jth transmit signal and the pth nonlinearity order, the signal model contains in fact all the different combinations of the exponents q_m and r_n, under the constraint that their sum is equal to p. This means that we can rewrite (7) as

$$x_j^{PA}(n) = \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{k} \sum_{m=0}^{M} h_{j,p,\mathbf{s}^k}(m) \times \prod_{q=1}^{N_T} x_q(n-m)^{s_q^k} x_q^*(n-m)^{s_{q+N_T}^k}, \quad \left\|\mathbf{s}^k\right\|_1 = p \tag{8}$$

where $\mathbf{s}^k$ is the kth combination of the $2N_t \times 1$ exponent vector $\mathbf{s}$, $h_{j,p,\mathbf{s}^k}(m)$ contains the corresponding coefficients, and $\|\cdot\|_1$ denotes the L^1-norm. Note that all the elements of $\mathbf{s}$ are non-negative integers, as per the signal model. To illustrate its structure, all the variations of $\mathbf{s}$ for $N_T = 1$ and $P = 3$ are written below:

$$\mathbf{s}^1 = \begin{bmatrix} 1 & 0 \end{bmatrix}^T,\ \mathbf{s}^2 = \begin{bmatrix} 0 & 1 \end{bmatrix}^T,\ \mathbf{s}^3 = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$$

$$\mathbf{s}^4 = \begin{bmatrix} 2 & 1 \end{bmatrix}^T,\ \mathbf{s}^5 = \begin{bmatrix} 3 & 0 \end{bmatrix}^T,\ \mathbf{s}^6 = \begin{bmatrix} 0 & 3 \end{bmatrix}^T$$

After the PAs, there is typically also some additional crosstalk between the transmitters, referred to as *linear crosstalk*. Taking also this phenomenon into account, the final output signal for the jth transmitter can be written as

$$\begin{aligned} x_j^{TX}(n) &= \sum_{l=1}^{N_T} \beta_{lj} x_l^{PA}(n) \\ &= \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{k} \sum_{m=0}^{M} \sum_{l=1}^{N_T} \beta_{lj} h_{l,p,\mathbf{s}^k}(m) \times \prod_{q=1}^{N_T} x_q(n-m)^{s_q^k} x_q^*(n-m)^{s_{q+N_T}^k}, \quad \left\|\mathbf{s}^k\right\|_1 = p \end{aligned} \tag{9}$$

where β_{lj} is the crosstalk coefficient between the lth and jth transmitters. It can be observed that the essential signal model remains the same as in (8), but with modified coefficients written as

$$\acute{h}_{j,p,\mathbf{s}^k}(m) = \sum_{l=1}^{N_T} \beta_{lj} h_{l,p,\mathbf{s}^k}(m). \tag{10}$$

Denoting the MIMO propagation channel impulse response from TX antenna j to RX antenna i by $c_{ij}(l)$, $l = 0, 1, \ldots, L$, the received SI signal at RX antenna i ($i = 1, 2, \ldots, N_R$) can now be written as

$$\begin{aligned} z_i(n) &= \sum_{j=1}^{N_T} \sum_{l=0}^{L} c_{ij}(l) x_j^{TX}(n-l) \\ &= \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{k} \sum_{j=1}^{N_T} \sum_{l=0}^{L} \sum_{m=0}^{M} c_{ij}(l) \acute{h}_{j,p,\mathbf{s}^k}(m) \times \prod_{q=1}^{N_T} x_q(n-m-l)^{s_q^k} x_q^*(n-m-l)^{s_{q+N_T}^k}, \quad \left\|\mathbf{s}^k\right\|_1 = p \\ &= \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{k} \sum_{m=0}^{M+L} \widetilde{h}_{i,p,\mathbf{s}^k}(m) \times \prod_{q=1}^{N_T} x_q(n-m)^{s_q^k} x_q^*(n-m)^{s_{q+N_T}^k}, \quad \left\|\mathbf{s}^k\right\|_1 = p \end{aligned} \tag{11}$$

Again, the signal model still remains the same as in (8), but with slightly modified coefficients, which are obtained from

$$\widetilde{h}_{i,p,\mathbf{s}^k}(m) = \sum_{j=1}^{N_T} \sum_{l=0}^{m} c_{ij}(l) \acute{h}_{j,p,\mathbf{s}^k}(m-l).$$

The new memory length of the received signal model is also increased from M to $M+L$. The input signal of the ith receiver ($z_i(n)$) is then further processed by the RF canceller and the actual receiver chain. Note that the above signal model in (11) also applies to circulator and electrical balance duplexer-based implementations, where each transmitter and receiver pair share the same antenna [32], and hence it is generic in that respect.

2.2 RF cancellation

To ensure an extensive analysis and derivation for the proposed digital cancellation algorithm, we consider three different RF cancellation solutions. The first technique is similar to what has been used, e.g, in [5, 6], and it involves directly tapping the transmitter outputs to obtain the reference signals for RF cancellation. This method is based on purely analog processing, as the whole cancellation procedure is performed in the RF domain. The two other considered methods are based on auxiliary TX chains, which are used to produce the RF cancellation signal from digital baseband samples [1, 38, 39]. We call this latter approach hybrid RF cancellation to distinguish it from purely analog cancellation. Furthermore, we consider both linear and nonlinear preprocessing to be used with this auxiliary transmitter based RF cancellation.

2.2.1 RF cancellation with transmitter output signals

In this RF cancellation method, the output of each TX chain is tapped, and subtracted from each of the received

signals after suitable gain, phase and delay adjustments. These RF cancellers can be either single-tap or multi-tap [9, 40], for which reason we denote them with impulse responses $h_{ij}^{RF}(l)$, operating on the TX output signals $x_j^{TX}(n)$. The coefficients are obviously chosen such that they model the MIMO coupling channel coefficients in $c_{ij}(n)$ as accurately as possible. The RF cancellation signal for the ith receiver can thus be written as

$$z_i^c(n) = \sum_{j=1}^{N_T} \sum_{l=0}^{L'} h_{ij}^{RF}(l) x_j^{TX}(n-l), \tag{12}$$

where L' is the number of taps in the RF canceller. It can be easily shown that the cancellation signal is of similar form as the actual received signal in (11), with coefficients of the form

$$\check{h}_{i,p,\mathbf{s}^k}^{RF}(m) = \sum_{j=1}^{N_T} \sum_{l=0}^{m} h_{ij}^{RF}(l) \acute{h}_{j,p,\mathbf{s}^k}(m-l)$$

and a memory length of $M + L'$. Thus, the received SI signal of receiver i, after this type of analog RF cancellation, becomes

$$\begin{aligned} r_i(n) &= z_i(n) - z_i^c(n) \\ &= \sum_{\substack{p=1 \\ p \text{ odd}}}^{P} \sum_{\left\|\mathbf{s}^k\right\|_1 = p}^{k} \sum_{m=0}^{M+\max(L,L')} \Big(\widetilde{h}_{i,p,\mathbf{s}^k}(m) \\ &\quad - \check{h}_{i,p,\mathbf{s}^k}^{RF}(m) \Big) \prod_{q=1}^{N_T} x_q(n-m)^{s_q^k} x_q^*(n-m)^{s_{q+N_T}^k}, \end{aligned} \tag{13}$$

Hence, the structure of the RF canceller output signal model is still of the same form as in (11), but with modified coefficients expressed as $\check{h}_{i,p,\mathbf{s}^k}(m) = \widetilde{h}_{i,p,\mathbf{s}^k}(m) - \check{h}_{i,p,\mathbf{s}^k}^{RF}(m)$.

This type of purely analog RF cancellation calls for $N_T \times N_R$ canceller circuits to be implemented in the device, one canceller from each transmitter to each receiver. The complexity may become prohibitive when the number of antennas is significantly increased and, thereby, when implementing a high order full-duplex MIMO device, alternative methods for RF cancellation might have to be considered.

2.2.2 Hybrid RF cancellation using auxiliary transmitters with linear preprocessing

One such alternative RF canceller structure is the hybrid method, which utilizes extra transmitter chains, one for each receiver, to upconvert and subtract estimated replicas of the SI signals from the received signals at RF [1, 38, 39]. In this case, linear MIMO filtering is already done at digital baseband on the transmit signals $x_j(n)$ with some estimated MIMO channel responses $h_{ij}^{RF}(l)$. Since the transmit signals from the different antennas can now be combined already in the digital domain, the analog hardware complexity of this type of an RF cancellation scheme scales with N_R instead of $N_T N_R$, and may prove to be more attractive with a high number of antennas. Note that in this subsection, we consider only linear processing for the hybrid RF canceller, and thereby IQ modulator imbalance or PA nonlinearity are not explicitly dealt with at this stage. The RF cancellation signal can in this case be written as

$$z_i^c(n) = \sum_{j=1}^{N_T} \sum_{l=0}^{L'} h_{ij}^{RF}(l) x_j(n-l), \tag{14}$$

which is a special case of the signal model in (11) with $P = 1$ and coefficients $\check{h}_{i,1,\mathbf{s}^k}^{RF}(m)$ consisting of $h_{ij}^{RF}(l)$ with proper $\mathbf{s}^k$. The signal after RF cancellation is again obtained as shown in (13), and with the final coefficients as

$$\begin{aligned} \check{h}_{i,p,\mathbf{s}^k}(m) &= \widetilde{h}_{i,p,\mathbf{s}^k}(m) - \check{h}_{i,p,\mathbf{s}^k}^{RF}(m),\ p = 1 \\ \check{h}_{i,p,\mathbf{s}^k}(m) &= \widetilde{h}_{i,p,\mathbf{s}^k}(m),\ p \geq 3 \end{aligned}$$

Also this model is essentially of the same form as (11), with the coefficients of the linear SI terms being affected by the hybrid RF cancellation procedure, while the other terms remain unchanged. This means that the observed SI signal in the receiver digital domain can still be modeled with the same signal model as in the case of pure analog RF cancellation (or no RF cancellation at all). Thus, from the perspective of the digital cancellation algorithm, it makes no difference whether RF cancellation is performed by tapping the transmitter output or by using auxiliary TX chains with linear preprocessing, although the RF cancellation performance itself might obviously be different for the considered methods.

2.2.3 Hybrid RF cancellation using auxiliary transmitters with nonlinear preprocessing

Yet another alternative RF cancellation technique utilizes auxiliary transmitters, but with nonlinear preprocessing, instead of purely linear processing. The estimated MIMO channel responses of the different nonlinear SI terms are now denoted by $h_{ij,p}^{RF}(l)$. In the forthcoming analysis, it is assumed that the auxiliary TX chains are linear. This is a relatively feasible assumption, since no PA is required due to the lower output power requirements. Now, the

cancellation signal obtained with this RF cancellation procedure can be expressed as

$$z_i^c(n) = \sum_{j=1}^{N_T} \sum_{\substack{p=1 \\ p\ odd}}^{P'} \sum_{l=0}^{L'} h_{ij,p}^{RF}(l) x_j(n-l)^{\frac{p+1}{2}} \times x_j^*(n-l)^{\frac{p-1}{2}}, \quad (15)$$

where P' is the nonlinearity order of the RF cancellation signals. Note that this signal model neglects IQ imbalance and crosstalk, since the RF canceller must only attenuate the SI such that the receiver is not saturated. Also this RF cancellation signal can be easily represented with a signal model of the same form as in (11). The coefficients $\check{h}_{i,p,\mathbf{s}^k}^{RF}(m)$ of the signal model now consist of $h_{ij,p}^{RF}(l)$ with the parameters p and $\mathbf{s}^k$ that correspond to the basis functions $x_j(n-l)^{\frac{p+1}{2}} x_j^*(n-l)^{\frac{p-1}{2}}$, and other coefficients are set to zero. Similar to the other RF cancellation schemes, after subtracting the cancellation signal from the received signal, as in (13), the signal model remains the same and its coefficients are $\check{h}_{i,p,\mathbf{s}^k}(m) = \widetilde{h}_{i,p,\mathbf{s}^k}(m) - \check{h}_{i,p,\mathbf{s}^k}^{RF}(m)$. Now, also some of the nonlinear SI terms are attenuated by RF cancellation, as they are modeled in the preprocessing stage.

Overall, it can be concluded that the essential structure of the observed SI signal in the digital domain is independent of the chosen method for RF cancellation. This means that, in the forthcoming analysis, the same digital cancellation algorithm can be applied in all the situations since the only difference between the three alternative RF cancellation schemes are the relative power levels of the various SI terms. However, as already mentioned, the RF cancellation performance is likely to differ between these techniques, and also the hardware and computational requirements are different for each RF canceller structure.

In the forthcoming analysis, we will refer to the parameters of the signal model in all cases by $\check{h}_{i,p,\mathbf{s}^k}(m)$, similar to the above derivations, even though the exact values of the different coefficients vary for different RF cancellation techniques. This notation will simplify the equations and make them more straightforward and illustrative. Hence, the signal after RF cancellation, which is then processed by the digital canceller, can be written as

$$r_i(n) = \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{\|\mathbf{s}^k\|_1 = p}^{k} \sum_{m=0}^{M+\max(L,L')} \check{h}_{i,p,\mathbf{s}^k}(m) \times \prod_{q=1}^{N_T} x_q(n-m)^{s_q^k} x_q^*(n-m)^{s_{q+N_T}^k}, \quad (16)$$

Note that this signal model implicitly incorporates also the IQ imbalance occurring in the receiver, even though it is omitted in the derivations for brevity [15].

2.3 Total number of basis functions in the overall model

In general, with the above cascaded modeling approach for IQ modulator and PA impairments with crosstalk between the transmitters, it can easily be shown that the total number of basis functions in (16) becomes

$$n_b = \sum_{\substack{p=1 \\ p\ odd}}^{P} \binom{p+2N_T-1}{2N_T-1}. \quad (17)$$

Figure 3 illustrates the number of basis functions for different nonlinearity orders and numbers of transmit antennas for the full signal model and also for the crosstalk-free signal model discussed below in Section 2.4. It is immediately obvious that with higher order MIMO systems, or with heavily nonlinear PAs, the number of basis functions becomes unacceptably high when utilizing the full signal model with crosstalk. Thus, it is necessary to determine methods that will decrease the number of basis functions, and thereby facilitate the estimation of the parameters of this signal model also in practice.

Luckily, many of the terms arising from the cascade of the impairments are so insignificant that they can be neglected with very little effect on the overall modeling accuracy. This will reduce the computational cost of such modeling and the corresponding cancellation procedure. In this work, we propose a specific preprocessing stage which can be used to decrease the dimensionality of the full signal model in (16). This is elaborated in more details in Section 3.2.

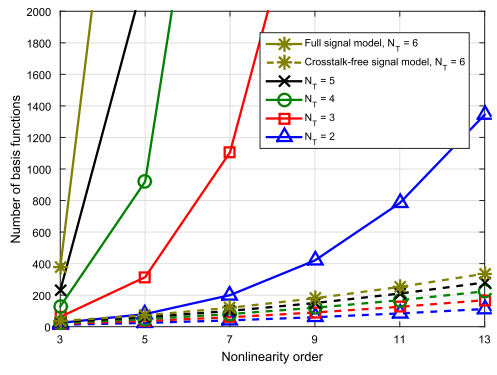


Fig. 3 The number of basis functions with respect to the nonlinearity order of the signal model (P). The curves have been plotted for different numbers of transmit antennas and for both the full signal model and the crosstalk-free signal model

2.4 Nonlinear signal model without crosstalk

Another simple way to decrease the number of basis functions is to neglect the crosstalk effect between the transmitters. Then, the cross terms between the different transmit signals will be removed, which obviously results in a significant decrease in the number of unknown parameters. Modifying (16) accordingly, we can write the signal model now as

$$r_i(n) = \sum_{j=1}^{N_T} \sum_{\substack{p=1 \\ p\ odd}}^{P} \sum_{q=0}^{p} \sum_{m=0}^{M+\max(L,L')} \breve{h}_{i,j,p,q}(m) \times x_j(n-m)^q x_j^*(n-m)^{p-q}, \tag{18}$$

where $\breve{h}_{i,j,p,q}(m)$ represents now the coupling channel corresponding to the considered SI signal terms propagating from the jth transmitter to the ith receiver. This signal model is also derived in [15], where it is briefly discussed and analyzed. For this reason, the detailed derivation process of (18) is omitted in this article.

Since now all the cross-terms are neglected from the signal model, the number of basis functions can be expressed as

$$n_b^{CT-free} = N_T \left(\frac{P+1}{2}\right)\left(\frac{P+1}{2}+1\right). \tag{19}$$

When investigating Fig. 3, it can be seen that this signal model results in a significant reduction of basis functions, when compared to the full signal model with crosstalk. With moderate crosstalk levels, it is therefore likely that using this signal model will provide a very favorable trade-off between cancellation performance and computational complexity. However, as already discussed, in highly integrated transceivers explicit modeling of the crosstalk between the transmitters is likely required in order to ensure sufficient cancellation performance [28].

3 Self-interference parameter estimation and digital cancellation

In this section, building on the previous modeling in, e.g., [15, 29], we will describe the proposed digital cancellation algorithm that models both IQ imbalance and PA nonlinearity in a MIMO full-duplex transceiver with crosstalk between the transmitters. In general, there are two possible approaches for nonlinear digital SI cancellation: (*i*) construct a linear-in-parameters model of the observed SI signal in the digital domain, including the different impairments, the MIMO propagation channel, and RF cancellation, estimate the unknown parameters of the model, and finally recreate and cancel the SI from the received signals; (*ii*) have separate models for the MIMO propagation channel and the transmitter impairments, estimate the unknown model parameters sequentially, and recreate and cancel the SI from the received signals. Typically the latter approach is computationally less demanding, but it requires a more elaborate estimation procedure. In this article, we consider the former approach, while the latter is left for future work.

3.1 Linear-in-parameters model

Having already derived a linear-in-parameters signal model in Section 2, presented in (16), the next step is to estimate its parameters in $\breve{h}_{i,p,\mathbf{s}^k}(m)$. After this, the estimated parameters are used to regenerate the SI signals, which are then subtracted from the received signals at digital baseband to obtain cancellation. Figure 4 shows the whole digital cancellation procedure on a fundamental level.

Denoting the desired signal of interest and additive noise at the ith receiver by $s_i(n)$ and $w_i(n)$, respectively, the overall received signal at digital baseband can be expressed as

$$y_i(n) = r_i(n) + s_i(n) + w_i(n). \tag{20}$$

The corresponding output of the digital SI canceller is then

$$\hat{s}_i(n) = y_i(n) - \hat{r}_i(n), \tag{21}$$

where $\hat{r}_i(n)$ denotes the SI estimate obtained using the signal model in (16) with estimated parameters, written as

$$\hat{r}_i(n) = \sum_{\substack{p=1 \\ p\ odd}}^{\bar{P}} \sum_{k \atop \|\mathbf{s}^k\|_1 = p} \sum_{m=-M_1}^{M_2} \hat{\breve{h}}_{i,p,\mathbf{s}^k}(m) \times \prod_{q=1}^{N_T} x_q(n-m)^{s_q^k} x_q^*(n-m)^{s_{q+N_T}^k}. \tag{22}$$

Here, $\bar{P}$ is the nonlinearity order of the digital canceller, M_1 is the number of pre-cursor taps, M_2 is the number of post-cursor taps, and $\hat{\breve{h}}_{i,p,\mathbf{s}^k}(m)$ contains the estimated parameters of the signal model. The pre-cursor taps are

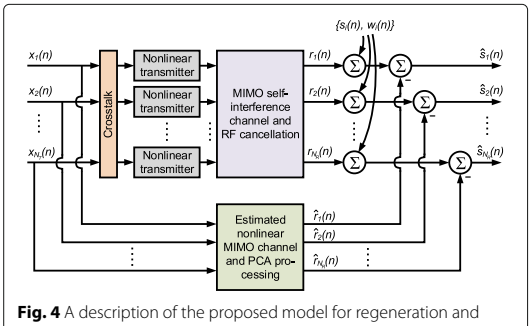


Fig. 4 A description of the proposed model for regeneration and cancellation of nonlinear self-interference

introduced to model all the memory effects produced by the transmitter and RF cancellation circuitry.

3.1.1 Least-squares-based estimator

In this work, the actual parameter learning is performed with the widely used least squares (LS) estimation. For brevity, the parameter learning and digital cancellation procedure is here outlined only for the ith receiver, since the procedure is identical for all the receivers.

In practice, calculating the LS estimate requires knowledge of (i) the original transmitted data signal, (ii) the predetermined signal model in (16), and (iii) the observed received signal $y_i(n)$. In the considered MIMO full-duplex device, all of these are obviously known by the digital canceller. Since the LS estimation is performed using a block of data, the vector/matrix representations of the relevant signals with N observed samples are first defined as

$$\begin{aligned} &\mathbf{y}_i = \mathbf{r}_i + \mathbf{s}_i + \mathbf{w}_i, \text{ with} \\ &\mathbf{y}_i = \begin{bmatrix} y_i(n) & y_i(n+1) & \cdots & y_i(n+N-1) \end{bmatrix}^{\mathrm{T}} \end{aligned} \tag{23}$$

and $\mathbf{r}_i, \mathbf{s}_i, \mathbf{w}_i$ are defined in the same manner as $\mathbf{y}_i$. The error vector is then defined as

$$\mathbf{e}_i = \mathbf{y}_i - \hat{\mathbf{r}}_i \tag{24}$$

where the nonlinear SI estimate is

$$\hat{\mathbf{r}}_i = \boldsymbol{\Psi} \hat{\check{\mathbf{h}}}_i. \tag{25}$$

Here, $\boldsymbol{\Psi}$ is a horizontal concatenation of the convolution matrices defined as follows:

$$\boldsymbol{\Psi}_{i,p,\mathbf{s}^k} = \begin{bmatrix} \psi_{i,p,\mathbf{s}^k}(n+M_1) & \psi_{i,p,\mathbf{s}^k}(n+M_1-1) & \cdots & \psi_{i,p,\mathbf{s}^k}(n-M_2) \\ \psi_{i,p,\mathbf{s}^k}(n+M_1+1) & \psi_{i,p,\mathbf{s}^k}(n+M_1) & \cdots & \psi_{i,p,\mathbf{s}^k}(n-M_2+1) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{i,p,\mathbf{s}^k}(n+M_1+N-1) & \psi_{i,p,\mathbf{s}^k}(n+M_1+N-2) & \cdots & \psi_{i,p,\mathbf{s}^k}(n-M_2+N-1) \end{bmatrix},$$

where

$$\psi_{i,p,\mathbf{s}^k}(n) = \prod_{q=1}^{N_T} x_q(n)^{\mathbf{s}^k_q} x_q^*(n)^{\mathbf{s}^k_{q+N_T}},$$

with $p = 1, 3, \ldots, \bar{P}$, and $\mathbf{s}^k$ is each combination for which $\left\|\mathbf{s}^k\right\|_1 = p$, similar to the sum limits shown in (16). Overall, the number of concatenated matrices is given by the total number of basis functions in (17), since this is the amount of different combinations of $\mathbf{s}^k$ for all the nonlinearity orders.

Alternatively, in the crosstalk-free model $\boldsymbol{\Psi}$ consists of the concatenation of the matrices defined as follows:

$$\boldsymbol{\Psi}_{j,p,q} = \begin{bmatrix} \psi_{j,p,q}(n+M_1) & \psi_{j,p,q}(n+M_1-1) & \cdots & \psi_{j,p,q}(n-M_2) \\ \psi_{j,p,q}(n+M_1+1) & \psi_{j,p,q}(n+M_1) & \cdots & \psi_{j,p,q}(n-M_2+1) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{j,p,q}(n+M_1+N-1) & \psi_{j,p,q}(n+M_1+N-2) & \cdots & \psi_{j,p,q}(n-M_2+N-1) \end{bmatrix},$$

where $\psi_{j,p,q}(n) = x_j(n)^q x_j^*(n)^{p-q}$, with $j = 1, 2, \ldots, N_T$, $p = 1, 3, \ldots, \bar{P}$, and $q = 0, 1, \ldots, p$.

An estimate of the parameter vector $\check{\mathbf{h}}_i$, denoted by $\hat{\check{\mathbf{h}}}_i$, is a vertical concatenation of the vectors

$$\hat{\check{\mathbf{h}}}_{i,p,\mathbf{s}^k} = \begin{bmatrix} \hat{\check{h}}_{i,p,\mathbf{s}^k}(-M_1) & \cdots & \hat{\check{h}}_{i,p,\mathbf{s}^k}(M_2) \end{bmatrix}^T \tag{26}$$

In the crosstalk-free model, the parameter vector consists of the concatenation of vectors

$$\hat{\check{\mathbf{h}}}_{i,j,p,q} = \begin{bmatrix} \hat{\check{h}}_{i,j,p,q}(-M_1) & \cdots & \hat{\check{h}}_{i,j,p,q}(M_2) \end{bmatrix}^T \tag{27}$$

The LS estimate of the parameter vector $\check{\mathbf{h}}_i$ is then found as the solution which minimizes the power of the error vector $\mathbf{e}_i$, as

$$\begin{aligned} \hat{\check{\mathbf{h}}}_i &= \arg\min_{\check{\mathbf{h}}_i} \|\mathbf{e}_i\|^2 = \arg\min_{\check{\mathbf{h}}_i} \left\| \mathbf{y}_i - \boldsymbol{\Psi} \check{\mathbf{h}}_i \right\|^2 \\ &= \left(\boldsymbol{\Psi}^H \boldsymbol{\Psi}\right)^{-1} \boldsymbol{\Psi}^H \mathbf{y}_i, \end{aligned} \tag{28}$$

assuming full column rank in $\boldsymbol{\Psi}$.

3.2 Computationally efficient estimation with principal component analysis

Another approach to simplify the estimation procedure is to retain the cross-terms, and instead determine which of them are actually significant in terms of the cancellation performance. In this analysis, principal component analysis (PCA) [41] is used to decrease the number of parameters to be estimated. The idea behind the PCA is to determine which of the terms have the highest variance, providing valuable information regarding the significance of the different basis functions. In practice, PCA results in a transformation matrix, with which the original data matrix is multiplied. The size of the transformation matrix can be chosen to provide the desired number of parameters for the final estimation procedure.

There are also various alternative solutions for model complexity reduction, such as compressed sampling (CS) based techniques. Nevertheless, in this work, we choose to use the PCA since it is a straight-forward method for the complexity reduction of the proposed signal model, while also providing nearly the same performance as CS when high modeling accuracy is required [42]. Experimenting with different complexity reduction methods is an important future work item for us.

The first step in obtaining the desired PCA transformation matrix is to determine the least squares channel estimate given in (28) using all the basis functions. This estimate should be calculated with the highest possible transmit power, since the nonlinear SI terms that are negligible with the highest power will also be negligible with any lower transmit power. Hence, this reveals the terms, which can be omitted under the whole considered transmit power range. If the transceiver in question has more than one receiver chain, the channel estimation can be done individually for all of them, after which the mean value of the estimates is calculated. This is done to avoid having separate transformation matrices for each receiver, resulting in a decreased amount of required data storage. The hereby obtained coefficient vector, which is denoted by $\hat{\tilde{\mathbf{h}}}_0$, is used as an initial channel estimate for the full set of basis functions.

The next step is to determine the relative strengths of the different terms present in the SI signal. Using the initial channel estimate, this can be done by multiplying the original data matrix with the obtained estimate. Then, we get

$$\mathbf{\Psi}_0 = \left(\mathbf{1}\hat{\tilde{\mathbf{h}}}_0^T \right) \times \mathbf{\Psi}, \tag{29}$$

where $\mathbf{1}$ is a column vector consisting of 1s, and $\times$ denotes element-wise multiplication between two matrices. The matrix $\mathbf{\Psi}_0$ now contains all the SI terms in its columns, each multiplied with the corresponding coefficient of the initial channel estimate.

As a starting point for the PCA, the singular value decomposition of the normalized data matrix can be expressed as

$$\mathbf{\Psi}_0 = \mathbf{U}\Sigma\mathbf{V}^H, \tag{30}$$

where $\mathbf{U}$ and $\mathbf{V}$ are the matrices containing the left and right singular vectors, respectively, while Σ is a diagonal matrix consisting of the corresponding singular values. In this analysis, it is assumed that the singular values are in decreasing order. To minimize the possible numerical issues upon the PCA transformation, the actual transformation matrix is obtained in its normalized form, which is given by

$$\mathbf{W} = \mathbf{V}\Sigma^{-1}. \tag{31}$$

To control the number of parameters, part of the columns of the obtained matrix $\mathbf{W}$ can then be omitted. Based on the earlier assumption regarding the ordering of the singular values, the columns of the transformation matrix represent the different parameters in the descending order of their significance. Thus, by starting to remove the columns from the right, the number of parameters can be decreased with minimal effect on the modeling accuracy. Thus, denoting the number of chosen parameters with u, we can write the final transformation matrix as

$$\widetilde{\mathbf{W}} = \left[\mathbf{w}_1 \ \mathbf{w}_2 \ \cdots \ \mathbf{w}_u \right], \tag{32}$$

where $\mathbf{w}_i$ is the ith column of the matrix $\mathbf{W}$. Finally, the reduced data matrix can be calculated as

$$\widetilde{\mathbf{\Psi}} = \mathbf{\Psi}\widetilde{\mathbf{W}}. \tag{33}$$

The hereby obtained data matrix is then used in the least squares estimation as a replacement for the original data matrix $\mathbf{\Psi}$. It should also be noted that when generating the actual digital cancellation signal, the cancellation data matrix must be transformed with the same matrix $\widetilde{\mathbf{W}}$, as the SI channel estimate is only valid in this transformed space.

An important aspect to point out is that the transformation matrix $\widetilde{\mathbf{W}}$ is calculated only once with the highest transmit power, after which it can be used with all transmit powers to reduce the number of basis functions. Namely, since the strengths of the nonlinearities are directly proportional to the transmit power, the SI terms that are negligibly weak with the highest transmit power are at least as weak with the lower transmit powers, which means that the same SI terms can be omitted also then. This is also proven by the waveform simulations, the results of which will be discussed in Section 4. However, should the SI channel change drastically at any point, then the matrix $\widetilde{\mathbf{W}}$ must be recalculated to ensure that no significant memory taps are neglected.

In general, perhaps the most crucial design problem in the context of the PCA is to determine the optimal number of parameters to be included in the final model. This can be most easily determined experimentally by reducing the number of parameters until the obtained cancellation performance starts to drop. Also, the singular values in Σ can be used to calculate the percentage of the variance accounted for by the included basis functions. We will address this issue more closely with the help of waveform simulations in Section 4.

4 Performance simulations and analysis

The evaluation of the proposed scheme is now done with realistic waveform simulations, utilizing a comprehensive inband full-duplex transceiver model. It incorporates all the relevant impairments, and thereby the SI waveform represents a real-world scenario rather well. Below, we describe the waveform simulator in detail, after which the results are shown. As an important future work item, we aim to evaluate the proposed scheme also with actual RF measurements to confirm the results obtained here with the simulations.

4.1 Simulation setup and parameters

The waveform simulations are performed with Matlab, where all the relevant aspects of the full-duplex

transceiver are modeled. These include the nonlinearity of the PAs, the crosstalk between the transmitters (both before and after the PA), the multipath SI channel, the imperfect RF cancellation, nonlinearity of the receiver, IQ imbalance, phase noise, and the quantization upon analog-to-digital conversion, while the DAC/ADC nonlinearities are omitted also from the simulator model since we have not observed them to be a significant factor in our earlier RF measurements [3, 43]. This means that the simulator model is rather comprehensive and can be expected to provide realistic results, although they must still be confirmed with real-life measurements. Note that, since the focus of this work is on SI cancellation, the signal of interest is not present in any of the simulations. The RF cancellation is performed in all the cases using the transmitter output signal, since the essential signal model is not affected by the RF cancellation procedure, as shown in Section 2.2. The used waveform is a 20 MHz LTE downlink signal, which utilizes OFDM with a 4-QAM constellation. When modeling the phase noise, a common local oscillator for all the transmitters and receivers is assumed, which is a feasible assumption for an inband full-duplex device. All the relevant parameters of the waveform simulator are listed in Table 1, while the used phase noise characteristics are shown in Fig. 5.

In the forthcoming results, five different digital cancellers are considered, and they are as follows:

Table 1 The relevant parameters of the waveform simulator

Parameter	Value
Bandwidth	20 MHz
Sampling frequency	122.88 MHz
Number of TX/RX antennas	2/2
PA gain	27 dB
PA IIP3	13 dBm
Level of TX crosstalk before the PAs	−10 dB/varied
Level of TX crosstalk after the PAs	−10 dB
Receiver noise floor	−96.9 dBm
Phase noise characteristics	See Fig. 5
Transmit power	25 dBm/varied
SI channel length	20 taps
Antenna attenuation	40 dB
RF cancellation	30 dB
IRR (TX/RX)	25 dB
ADC bits	12
Parameter estimation sample size (N)	30,000/varied
Parameter estimation sample size for PCA	10,000
Nonlinearity order of the canceller (P)	5
Number of pre-cursor taps (M_1)	10
Number of post-cursor taps (M_2)	20

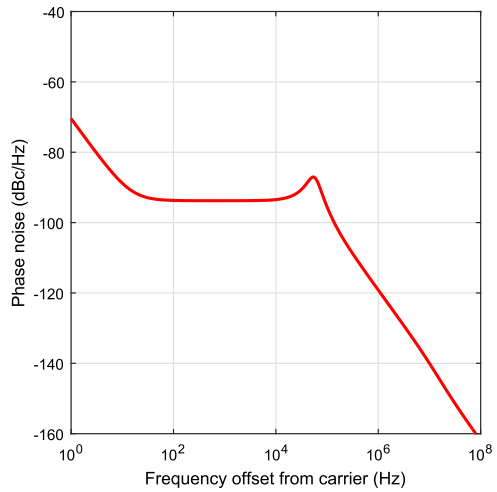


Fig. 5 The phase noise characteristics used in the waveform simulator, corresponding to a charge-pump type phase locked loop-based oscillator, taken from: http://datasheets.maximintegrated.com/en/ds/MAX2870.pdf

- Digital canceller with the full signal model in (22), including PCA processing to decrease the dimensionality and computational complexity
- Digital canceller with the full signal model in (22), but without any dimensionality reduction
- Digital canceller utilizing the N-input memory model from [26], which considers the nonlinearity of the PA and both linear and nonlinear crosstalk.
- Digital canceller with the crosstalk-free signal model in (18), from [15], where both the nonlinearity of the PA and the IQ imbalance are modeled.
- Digital canceller with a traditional linear signal model, where $P = 1$.

In all the cases, the same parameter estimation sample size is used for the different cancellers with $M_1 = 10$ and $M_2 = 20$ to ensure a fair comparison. The PCA matrix is calculated using 10 000 samples in the initial channel estimation stage. Furthermore, to avoid overfitting when estimating and cancelling the SI, separate portions of the signal are used for calculating the SI channel estimate and evaluating the actual SI cancellation performance.

4.2 Results

First, the signal spectra after the different digital cancellers are shown in Fig. 6 using the default parameters, alongside with the spectra of the RF cancelled signal and the receiver noise floor. It can be observed that only the digital cancellers utilizing the full signal model can obtain

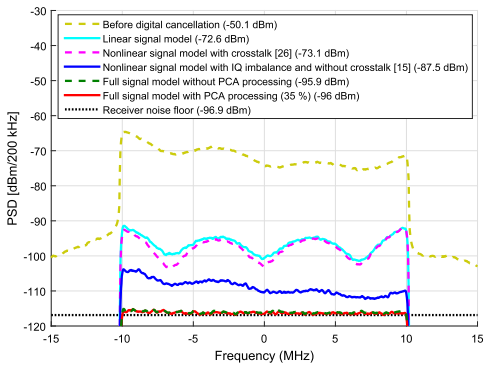


Fig. 6 The signal spectra after the different digital cancellers, alongside with the spectra of the RF cancelled signal and the receiver noise floor

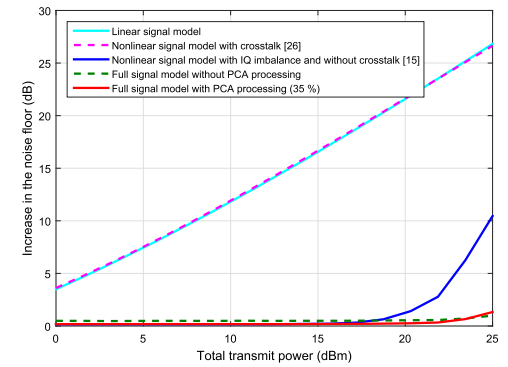


Fig. 7 The increase in the noise floor due to residual SI, with respect to the total transmit power

sufficient levels of SI cancellation. In particular, the digital canceller utilizing the linear signal model and the nonlinear crosstalk signal model from [26] perform very poorly since in this case IQ imbalance is the dominant source of distortion. The signal model from [15], on the other hand, has insufficient modeling accuracy since it does not take into account the crosstalk. Thereby, it is necessary to model both the IQ imbalance and the crosstalk, together with the nonlinearity of the PA, to obtain sufficient levels of digital cancellation. Furthermore, based on Fig. 6, the number of basis functions can be reduced to 35% without any reduction in the cancellation performance when using the full signal model.

Note that in this case the phase noise has no significant effect on the residual SI power since a common local oscillator between the transmitters and receivers is assumed. This results in a certain level of self-cancellation of the phase noise upon downconversion, considerably reducing its significance [44].

Figure 7 shows then the increase in the effective noise floor due to the residual SI for the different digital cancellers, with respect to the total transmit power. In other words, the closer to 0 dB the canceller achieves, the better is its overall SI cancellation performance. As expected, the linear canceller is not capable of efficient cancellation even with the lowest transmit powers, whereas the nonlinear cancellers with IQ imbalance modeling suppress the SI nearly perfectly up to transmit powers of 20 dBm. Moreover, the digital canceller utilizing the nonlinear crosstalk signal model from [26] performs very poorly with the whole transmit power range since it does not model the IQ imbalance, as already discussed.

With transmit powers beyond 20 dBm, the crosstalk effects begin to decrease also the accuracy of the crosstalk-free nonlinear signal model from [15]. On the other hand, the full signal models perform relatively well even with the highest transmit powers, resulting in only a very minor increase in the noise floor. Furthermore, as observed earlier, retaining only 35% of the terms after the PCA processing does not seem to decrease the accuracy of the signal model when compared to the full signal model with all the terms included. In fact, the performance of the digital canceller with the lower transmit powers is slightly improved by the dimensionality reduction since the smaller number of parameters results in a more accurate parameter vector estimate, and hence in more efficient cancellation.

To investigate the PCA-based dimensionality reduction in greater detail, Fig. 8 shows the increase in the noise floor with respect to the percentage of the terms included after the PCA, when using the full signal model in (22). The performance of the case without any PCA processing is also shown for reference. It can be observed from the

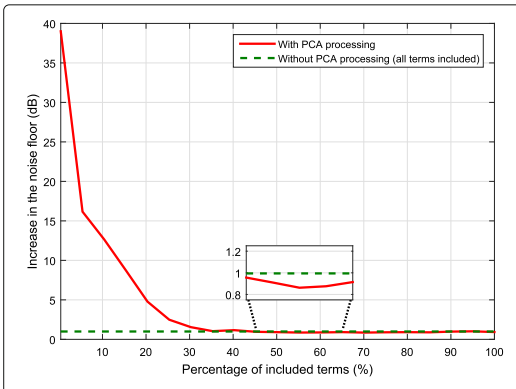


Fig. 8 The increase in the noise floor due to residual SI, with respect to the percentage of included terms

figure that there is a wide range of values for the percentage of included terms that provide the same cancellation performance. However, if the percentage of included terms goes significantly below 35%, the performance of the PCA-based canceller is rather poor. This is caused by the decreased accuracy of the signal model due to excluding some of the significant terms. Also note that when 50–80% of the terms are included, the PCA-based solution achieves slightly higher levels of SI cancellation than the canceller without PCA processing. The reason for this is the decreased variance of the parameter estimate, thanks to the smaller number of terms.

In order to minimize the computational complexity of the cancellation procedure, the number of included terms must obviously be minimized. Hence, the smallest number of terms that still provides the required performance is in this sense the optimal choice. Figure 8 indicates that, with the parameters considered in these simulations, the optimal percentage of included terms is roughly 35%, which corresponds to 840 coefficients with the considered nonlinearity order and number of memory taps.

Since the level of the crosstalk occurring before the transmitter PAs is obviously the most significant aspect in determining whether the full signal model is actually necessary, Fig. 9 shows then the performance of the different digital cancellers with different crosstalk levels. It can be observed that, with the considered transmit power of 25 dBm, the crosstalk has a rather significant effect already at the level of -20 dB, since using the nonlinear signal model without any crosstalk modeling from [15] results in a 3 dB higher noise floor than when using the full signal models. With higher crosstalk levels, the performance difference is obviously further emphasized. Furthermore, similar to the earlier observations, the signal models that do not model the IQ imbalance perform very poorly since it is the dominant source of distortion.

It can also be observed from Fig. 9 that a larger number of terms is required with the very high crosstalk levels. In particular, having only 35% of the terms retained results in a somewhat higher residual SI power than retaining all of the terms. This is explained by the fact that higher crosstalk levels also result in a larger number of significantly powerful SI terms. Nevertheless, the cancellation performance differences between the full signal models, with or without PCA processing, are still relatively small with these reasonable crosstalk levels.

In order to further investigate the differences in the computational complexity of the different digital cancellers, Fig. 10 shows their performance for different parameter estimation sample sizes (N). It can be observed that the signal models without sufficient modeling accuracy are not bottlenecked by the amount of available learning data, since their performance is largely unaffected by the value of N. The benefits of the PCA-based dimensionality reduction for the full signal model are also clearly apparent, since the case with 35% of the terms retained performs relatively well even with very small parameter estimation sample sizes. As opposed to this, without any dimensionality reduction, roughly $N = 24\,000$ is required to obtain a sufficiently accurate estimate of the parameters. Overall, it is hence clear that the PCA processing helps in significantly reducing the computational complexity of the digital SI cancellation procedure when utilizing the full signal model.

5 Conclusions

In this paper, a novel digital self-interference canceller for a nonlinear MIMO inband full-duplex transceiver was presented. The canceller is based on a comprehensive signal model for the SI observed in the digital domain, which includes the effect of crosstalk occurring between the transmit chains, while also incorporating the most

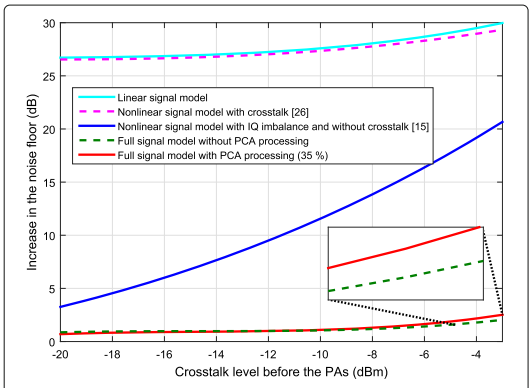


Fig. 9 The increase in the noise floor due to residual SI, with respect to the level of the crosstalk before the PAs

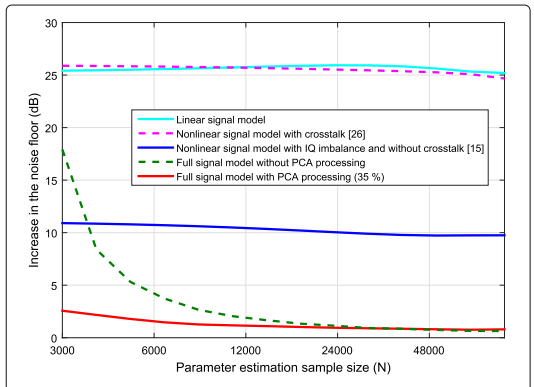


Fig. 10 The increase in the noise floor due to residual SI, with respect to the parameter estimation sample size (N)

significant RF imperfections. Furthermore, it was also shown that the signal model is valid for various different RF cancellers. To control the complexity of the cancellation procedure, a novel principal component analysis based scheme was then proposed, which can be used to control the number of parameters in the signal model. With the help of waveform simulations, the proposed digital canceller was shown to cancel the SI nearly perfectly, even when its computational complexity was significantly reduced using principal component analysis.

Appendix: Power amplifier output signal under crosstalk

Let us define a signal $y(n)$ as follows:

$$y(n) = \sum_{i=1}^{N} \alpha_i x_i(n),$$

where α_i is a scaling constant and $x_i(n)$ are known signals. To express an arbitrary integer power p of the signal $y(n)$ in terms of the signals $x_i(n)$, let us expand the corresponding equation accordingly.

$$y(n)^p = \left(\alpha_1 x_1(n) + \sum_{i=2}^{N} \alpha_i x_i(n)\right)^p.$$

Applying now the binomial theorem to the above expression, we obtain

$$y(n)^p = \sum_{k_1=0}^{p} \binom{p}{k_1} (\alpha_1 x_1(n))^{k_1} \left(\sum_{i=2}^{N} \alpha_i x_i(n)\right)^{p-k_1}$$

Applying the binomial theorem in a similar manner to the expression $\left(\sum_{i=2}^{N} \alpha_i x_i(n)\right)^{p-k_1}$, we get

$$\begin{aligned} y(n)^p &= \sum_{k_1=0}^{p} \left[\binom{p}{k_1} (\alpha_1 x_1(n))^{k_1} \right. \\ &\quad \left. \times \left(\alpha_2 x_2(n) + \sum_{i=3}^{N} \alpha_i x_i(n)\right)^{p-k_1}\right] \\ &= \sum_{k_1=0}^{p} \left[\binom{p}{k_1} (\alpha_1 x_1(n))^{k_1} \left[\sum_{k_2=0}^{p-k_1} \left[\binom{p-k_1}{k_2}\right.\right.\right. \\ &\quad \left.\left.\left. \times (\alpha_2 x_2(n))^{k_2} \left(\sum_{i=3}^{N} \alpha_i x_i(n)\right)^{p-k_1-k_2}\right]\right]\right] \\ &= \sum_{k_1=0}^{p} \sum_{k_2=0}^{p-k_1} \binom{p}{k_1}\binom{p-k_1}{k_2} (\alpha_1 x_1(n))^{k_1} \\ &\quad \times (\alpha_2 x_2(n))^{k_2} \left(\sum_{i=3}^{N} \alpha_i x_i(n)\right)^{p-k_1-k_2}. \end{aligned}$$

Applying the binomial theorem again to the expression $\left(\sum_{i=3}^{N} \alpha_i x_i(n)\right)^{p-k_1-k_2}$ and continuing in a similar manner, we finally obtain the following equation:

$$\begin{aligned} \left(\sum_{i=1}^{N} \alpha_i x_i(n)\right)^p &= \sum_{k_1=0}^{p} \sum_{k_2=0}^{p-k_1} \cdots \sum_{k_{N-1}=0}^{p-k_1-\cdots-k_{N-2}} \binom{p}{k_1} \\ &\quad \times \binom{p-k_1}{k_2} \cdots \binom{p-k_1-\cdots-k_{N-2}}{k_{N-1}} \\ &\quad \times x_1(n)^{k_1} x_2(n)^{k_2} \cdots x_N(n)^{p-k_1-\cdots-k_{N-1}} \\ &= \sum_{k_1=0}^{p} \sum_{k_2=0}^{p-k_1} \cdots \sum_{k_{N-1}=0}^{p-k_1-\cdots-k_{N-2}} A_{k_1,\ldots,k_{N-1}} \\ &\quad \times x_1(n)^{k_1} x_2(n)^{k_2} \cdots x_N(n)^{p-k_1-\cdots-k_{N-1}}, \end{aligned}$$

where $A_{k_1,\ldots,k_{N-1}}$ is a constant.

Acknowledgements

The research work leading to these results was funded by the Academy of Finland (under the projects #259915, #301820, and #304147), the Finnish Funding Agency for Technology and Innovation (Tekes, under the TAKE-5 project), Tampere University of Technology Graduate School, Nokia Foundation, Tuula and Yrjö Neuvo Research Fund, and Emil Aaltonen Foundation.

Competing interests

The authors declare that they have no competing interests.

Received: 30 June 2016 Accepted: 3 January 2017
Published online: 02 February 2017

References

1. M Duarte, C Dick, A Sabharwal, Experiment-driven characterization of full-duplex wireless systems. IEEE Trans. Wirel. Commun. **11**(12), 4296–4307 (2012)
2. M Heino, D Korpi, T Huusari, E Antonio-Rodríguez, S Venkatasubramanian, T Riihonen, L Anttila, C Icheln, K Haneda, R Wichman, M Valkama, Recent advances in antenna design and interference cancellation algorithms for in-band full-duplex relays. IEEE Commun. Mag. **53**(5), 91–101 (2015)
3. D Korpi, Y-S Choi, T Huusari, S Anttila, L Talwar, M Valkama, in *Proc. IEEE Global Communications Conference (GLOBECOM)*. Adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio: algorithms and RF measurements (Institute of Electrical and Electronics Engineers (IEEE), New York, 2015)
4. T Riihonen, S Werner, R Wichman, Mitigation of loopback self-interference in full-duplex MIMO relays. IEEE Trans. Signal Process. **59**(12), 5983–5993 (2011)
5. JI Choi, M Jain, K Srinivasan, P Levis, S Katti, in *Proc. 16th Annual International Conference on Mobile Computing and Networking*. Achieving single channel full duplex wireless communication (Association for Computing Machinery (ACM), New York, 2010), pp. 1–12
6. M Jain, JI Choi, T Kim, D Bharadia, S Seth, K Srinivasan, P Levis, S Katti, P Sinha, in *Proc. 17th Annual International Conference on Mobile Computing and Networking*. Practical, real-time, full duplex wireless (Association for Computing Machinery (ACM), New York, 2011), pp. 301–312
7. D Bharadia, S Katti, in *Proc. 11th USENIX Conference on Networked Systems Design and Implementation*. Full duplex MIMO radios (USENIX, the Advanced Computing Systems Association, Berkeley, 2014), pp. 359–372
8. A Sabharwal, P Schniter, D Guo, D Bliss, S Rangarajan, R Wichman, In-band full-duplex wireless: Challenges and opportunities. IEEE J. Sel. Areas Commun. **32**(9), 1637–1652 (2014)

9. Y-S Choi, H Shirani-Mehr, Simultaneous transmission and reception: algorithm, design and system level performance. IEEE Trans. Wirel. Commun. **12**(12), 5992–6010 (2013)
10. E Everett, A Sahai, A Sabharwal, Passive self-interference suppression for full-duplex infrastructure nodes. IEEE Trans. Wirel. Commun. **13**(2), 680–694 (2014)
11. G Liu, FR Yu, H Ji, VCM Leung, X Li, In-band full-duplex relaying: a survey, research issues and challenges. IEEE Commun. Surv. Tutorials. **17**(2), 500–524 (2015)
12. Z Zhang, Y Shen, S Shao, W Pan, Y Tang, in *Proc. Sixth International Conference on Wireless Communications and Signal Processing (WCSP)*. Full duplex 2x2 MIMO radios (Institute of Electrical and Electronics Engineers (IEEE), New York, 2014), pp. 1–6
13. A Balatsoukas-Stimming, P Belanovic, K Alexandris, A Burg, in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*. On self-interference suppression methods for low-complexity full-duplex MIMO (Institute of Electrical and Electronics Engineers (IEEE), New York, 2013), pp. 992–997
14. K Alexandris, A Balatsoukas-Stimming, A Burg, in *Proc. IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. Measurement-based characterization of residual self-interference on a full-duplex MIMO testbed (Institute of Electrical and Electronics Engineers (IEEE), New York, 2014), pp. 329–332
15. L Anttila, D Korpi, E Antonio-Rodríguez, R Wichman, M Valkama, in *Proc. IEEE Globecom Workshops*. Modeling and efficient cancellation of nonlinear self-interference in MIMO full-duplex transceivers (Institute of Electrical and Electronics Engineers (IEEE), New York, 2014), pp. 862–868
16. BP Day, AR Margetts, DW Bliss, P Schniter, Full-duplex bidirectional MIMO: achievable rates under limited dynamic range. IEEE Trans. Signal Process. **60**(7), 3702–3713 (2012)
17. Z Zhang, Z Chen, M Shen, B Xia, Spectral and energy efficiency of multipair two-way full-duplex relay systems with massive MIMO. IEEE J. Sel. Areas Commun. **34**(4), 848–863 (2016)
18. X Xia, Y Xu, K Xu, D Zhang, W Ma, Full-duplex massive MIMO AF relaying with semiblind gain control. IEEE Trans. Veh. Technol. **65**(7), 5797–5804 (2016)
19. X Xia, D Zhang, K Xu, W Ma, Y Xu, Hardware impairments aware transceiver for full-duplex massive MIMO relaying. IEEE Trans. Signal Process. **63**(24), 6565–6580 (2015)
20. T Huusari, Y-S Choi, P Liikkanen, D Korpi, S Talwar, M Valkama, in *Proc. IEEE 81st Vehicular Technology Conference (VTC Spring)*. Wideband self-adaptive RF cancellation circuit for full-duplex radio: operating principle and measurements (Institute of Electrical and Electronics Engineers (IEEE), New York, 2015)
21. HQ Ngo, HA Suraweera, M Matthaiou, EG Larsson, Multipair full-duplex relaying with massive arrays and linear processing. IEEE J. Sel. Areas Commun. **32**(9), 1721–1737 (2014)
22. R Zayani, R Bouallegue, D Roviras, in *Proc. 21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*. Crossover neural network predistorter for the compensation of crosstalk and nonlinearity in MIMO OFDM systems (Institute of Electrical and Electronics Engineers (IEEE), New York, 2010), pp. 966–970
23. S Amin, PN Landin, P Händel, D Rönnow, Behavioral modeling and linearization of crosstalk and memory effects in RF MIMO transmitters. IEEE Trans. Microw. Theory Tech. **62**(4), 810–823 (2014)
24. T Sadeghpour, RA Alhameed, NT Ali, ITE Elfergani, Y Dama, OO Anoh, in *Proc. 18th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. Linear and nonlinear crosstalk in MIMO OFDM transceivers (Institute of Electrical and Electronics Engineers (IEEE), New York, 2011), pp. 504–507
25. J Qi, S Aïssa, in *Proc. IEEE Wireless Communications and Networking Conference*. Analysis and compensation for the joint effects of HPA nonlinearity, I/Q imbalance and crosstalk in MIMO beamforming systems (Institute of Electrical and Electronics Engineers (IEEE), New York, 2011), pp. 1562–1567
26. D Saffar, N Boulejfen, FM Ghannouchi, A Gharsallah, M Helaoui, Behavioral modeling of MIMO nonlinear systems with multivariable polynomials. IEEE Trans. Microw. Theory Tech. **59**(11), 2994–3003 (2011)
27. MV Amiri, SA Bassam, M Helaoui, FM Ghannouchi, in *Proc. 15th IEEE International Workshop on Computer Aided Modeling, Analysis and Design of Communication Links and Networks (CAMAD)*. Matrix-based orthogonal polynomials for MIMO transmitter linearization (Institute of Electrical and Electronics Engineers (IEEE), New York, 2010), pp. 57–60
28. A Abdelhafiz, L Behjat, FM Ghannouchi, M Helaoui, O Hammi, A high-performance complexity reduced behavioral model and digital predistorter for MIMO systems with crosstalk. IEEE Trans. Commun. **64**(5), 1996–2004 (2016)
29. L Anttila, D Korpi, Syrjälä, M Valkama, in *Proc. 47th Asilomar Conference on Signals, Systems and Computers*. Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers (Institute of Electrical and Electronics Engineers (IEEE), New York, 2013), pp. 1193–1198
30. D Korpi, L Anttila, Syrjälä, M Valkama, Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver. IEEE J. Sel. Areas Commun. **32**(9), 1674–1687 (2014)
31. A Balatsoukas-Stimming, A Austin, P Belanovic, A Burg, Baseband and RF hardware impairments in full-duplex wireless systems: experimental characterisation and suppression. EURASIP J. Wireless Commun. Netw. **2015**(142), 142 (2015)
32. D Korpi, T Huusari, Y-S Choi, L Anttila, S Talwar, M Valkama, in *Proc. IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. Digital self-interference cancellation under nonideal RF components: advanced algorithms and measured performance (Institute of Electrical and Electronics Engineers (IEEE), New York, 2015), pp. 286–290
33. L Anttila, *Digital front-end signal processing with widely-linear signal models in radio devices*, PhD thesis, Tampere University of Technology. (Tampere University of Technology, Tampere, 2011). http://urn.fi/URN:ISBN:978-952-15-2978-8
34. M Isaksson, D Wisell, D Ronnow, A comparative analysis of behavioral models for RF power amplifiers. IEEE Trans. Microw. Theory Technol. **54**(1), 348–359 (2006)
35. L Ding, GT Zhou, DR Morgan, Z Ma, JS Kenney, J Kim, CR Giardina, A robust digital baseband predistorter constructed using memory polynomials. IEEE Trans. Commun. **52**(1), 159–165 (2004)
36. L Anttila, P Handel, M Valkama, Joint mitigation of power amplifier and I/Q modulator impairments in broadband direct-conversion transmitters. IEEE Trans. Microw. Theory Technol. **58**(4), 730–739 (2010)
37. H Ku, JS Kenney, Behavioral modeling of nonlinear RF power amplifiers considering memory effects. IEEE Trans. Microw. Theory Technol. **51**(12), 2495–2504 (2003)
38. M Duarte, A Sabharwal, in *Proc. 44th Asilomar Conference on Signals, Systems, and Computers*. Full-duplex wireless communications using off-the-shelf radios: feasibility and first results (Institute of Electrical and Electronics Engineers (IEEE), New York, 2010), pp. 1558–1562
39. M Duarte, A Sabharwal, V Aggarwal, R Jana, KK Ramakrishnan, CW Rice, NK Shankaranarayanan, Design and characterization of a full-duplex multiantenna system for WiFi networks. IEEE Trans. Veh. Technol. **63**(3), 1160–1177 (2014)
40. D Bharadia, E McMilin, S Katti, in *Proc. SIGCOMM'13*. Full duplex radios (Association for Computing Machinery (ACM), New York, 2013), pp. 375–386
41. IT Jolliffe, *Principal component analysis, 2nd Ed.* (Springer, New York, 2002)
42. A Abdelhafiz, A Kwan, O Hammi, FM Ghannouchi, Digital predistortion of LTE-A power amplifiers using compressed-sampling-based unstructured pruning of volterra series. IEEE Trans. Microw. Theory Technol. **62**(11), 2583–2593 (2014)
43. D Korpi, J Tamminen, M Turunen, T Huusari, Y-S Choi, L Anttila, S Talwar, M Valkama, Full-duplex mobile device: pushing the limits. IEEE Commun. Mag. **54**(9), 80–87 (2016)
44. Syrjälä, K Yamamoto, M Valkama, Analysis and design specifications for full-duplex radio transceivers under RF oscillator phase-noise with arbitrary spectral shape. IEEE Trans. Veh. Technol. **65**(8), 6782–6788 (2016)

Publication 7

D. Korpi, T. Riihonen, A. Sabharwal, and M. Valkama, "Transmit power optimization and feasibility analysis of self-backhauling full-duplex radio access systems," *IEEE Transactions on Wireless Communications*, under review after revision, 2017.

Transmit Power Optimization and Feasibility Analysis of Self-backhauling Full-Duplex Radio Access Systems

Dani Korpi, Taneli Riihonen, Ashutosh Sabharwal, and Mikko Valkama

***Abstract*—We analyze an inband full-duplex access node that is serving mobile users while simultaneously connecting to a core network over a wireless backhaul link, utilizing the same frequency band for all communication tasks.We analyze an inband full-duplex access node that is serving mobile users while simultaneously connecting to a core network over a wireless backhaul link, utilizing the same frequency band for all communication tasks. Such wireless self-backhauling is an intriguing option for the next generation wireless systems since a wired backhaul connection might not be economically viable if the access nodes are deployed densely. In particular, we derive the optimal transmit power allocation for such a system in closed form under Quality-of-Service (QoS) requirements, which are defined in terms of the minimum data rates for each mobile user. For comparison, the optimal transmit power allocation is solved also for two reference scenarios: a purely half-duplex access node, and a relay-type full-duplex access node. Based on the obtained expressions for the optimal transmit powers, we then show that the systems utilizing a full-duplex capable access node have a fundamental feasibility boundary, meaning that there are circumstances under which the QoS requirements cannot be fulfilled using finite transmit powers. This fundamental feasibility boundary is also derived in closed form. The feasibility boundaries and optimal transmit powers are then numerically evaluated in order to compare the different communication schemes. In general, utilizing the purely full-duplex access node results in the lowest transmit powers for all the communicating parties, although there are some network geometries under which such a system is not capable of reaching the required minimum data rates. In addition, the numerical results indicate that a full-duplex capable access node is best suited for relatively small cells.**

***Index Terms*—Self-backhauling, full-duplex wireless, massive MIMO, transmit power optimization.**

I. Introduction

WIRELESS inband full-duplex communications is widely considered to be one of the key enabling technologies in achieving the required throughput gains of the future 5G networks. Its basic idea is to allow a radio to transmit and receive data signals simultaneously using the same center frequency, and hence it has the capability to double the spectral efficiency of the existing systems as long as its full potential can be harnessed properly [1]–[5]. Many real-world demonstrations of inband full-duplex radios have already been developed by various research groups, which indicates that the concept is indeed feasible [1], [5]–[8]. In addition, the framework and theoretical boundaries of inband full-duplex radios have been studied extensively in the recent years [2], [9]–[15].

In terms of a practical implementation, the fundamental issue for inband full-duplex devices is the coupling of the own transmit signal to the receiver. In particular, since the transmission and reception occur simultaneously over the same frequency band, the transceiver will inherently receive its own transmit signal. What makes this especially problematic is the extremely high power level of the own transmission at this stage, which means that it will completely drown out the intended received signal. This phenomenon is typically referred to as self-interference (SI), and reducing its effect has been one of the main research areas in this field. The various proposed SI cancellation solutions [7], [16]–[20] and actual implementations and measurements already show that solving the problem of SI is not far from reality [1], [6]–[8], [20], [21].

In addition to SI cancellation, a large portion of the research has also focused on how to best make use of the full-duplex capability on a network level [4], [22]–[24]. This is a tedious issue since in many applications the traffic requirements are highly asymmetric between the two communication directions, such as in mobile networks [25]. Because the inband full-duplex principle requires completely symmetric traffic to realize the doubling of spectral efficiency at radio link level, this will compromise the potential throughput gains it can provide in practice. Thus, more advanced schemes are likely needed in order to realize the full potential of inband full-duplex radios in practical network scenarios. One such option is employing a full-duplex access node (AN) in an otherwise legacy half-duplex mobile cell [4], [23], [24], [26], thereby allowing the AN to simultaneously serve the uplink (UL) and the downlink (DL) using the very same frequency resources. With proper multiplexing and active scheduling, such a scheme enables the AN to fully exploit its full-duplex capability in both directions [4].

In this paper, the above type of a scheme will be analyzed and developed further under a scenario where installing wired backhaul links for all the cells is not feasible. This means that wireless *self-backhauling*, where the same frequency band is also used to backhaul the UL and DL data, is required [24],

D. Korpi, T. Riihonen, and M. Valkama are with the Laboratory of Electronics and Communications Engineering, Tampere University of Technology, Tampere, Finland, e-mail: dani.korpi@tut.fi.

A. Sabharwal is with the Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005, USA.

The research work leading to these results was funded by the Academy of Finland (under the projects #301820, #304147, and #310991), the Finnish Funding Agency for Technology and Innovation (Tekes, under the TAKE-5 project), Tampere University of Technology Graduate School, Nokia Foundation, Tuula and Yrjö Neuvo Research Fund, and Emil Aaltonen Foundation.

[27]–[33]. This type of a situation can occur, for instance, due to densely deployed cells, a probable scenario in the future 5G networks [34], [35]. Hence, in addition to communicating with the user equipments (UEs), also the backhaul data is transferred inband with a wireless point-to-point link between the AN and a so-called backhaul node (BN). The BN then further connects to the actual core network using either a wired or a wireless link. As the self-backhauling is performed on the same frequency band as the DL and UL data transfer, no additional spectral resources are needed, which further improves the applicability of such a solution.

This type of inband self-backhauling has been also investigated in the earlier literature. Therein, most works have considered a relay-type AN that is directly forwarding the signals transmitted by the UL UEs to the BN, or vice versa [27]–[30], [32], [36], [37]. The reason for this is likely the fact that such a relay-type AN is more or less directly compatible with the existing networks, as it would essentially just extend the range of the BN or macro base station (BS). In particular, in [27], [36], the power control of such a relay-type AN is investigated, and the performance of both half-duplex and full-duplex operation modes is then compared. The findings obtained in [27], where the transmit powers are numerically optimized, indicate that the full-duplex AN can obtain higher throughputs than the corresponding half-duplex system. The same conclusion is reached in [36], where the spectral efficiency of a similar system is maximized by solving the optimal power allocation for both full-duplex and half-duplex ANs, with the power allocation of the former being solved using an iterative algorithm.

Moreover, the effect of radio resource management (RRM) on the performance of the relay-type full-duplex AN is investigated in [28], where the resulting solution is shown to outperform the half-duplex benchmark scheme. The work in [29], on the other hand, investigates different beamforming solutions for a BN with massive antenna arrays, although no full-duplex operation is assumed in any of the nodes therein. The DL coverage of a relay-type self-backhauling AN is then analyzed in [30], [37]. The findings in [37] indicate that, while the throughput of the network with full-duplex-capable ANs is almost doubled in comparison to the half-duplex systems, the increased interference levels result in a somewhat smaller coverage. The results obtained in [30] suggest, on the other hand, that on a network level it may be better to have also some ANs that perform the self-backhauling on a different frequency band, in order to reduce the interference levels. The work in [32] analyzes the throughput and outage probability of a relay-type full-duplex AN under an antenna selection scheme where individual transmit and receive antennas are chosen in the AN based on a given criterion. Again, the full-duplex AN is shown to usually outperform the corresponding half-duplex AN.

All in all, even though different inband self-backhauling solutions for small cells have been investigated in the earlier literature, none of the above works have considered a scenario where also the UL and DL transmissions are performed simultaneously on the same center frequency. Considering the promising findings regarding the relay-type scenario where the DL and UL are separated either in time or in frequency, this means that the purely full-duplex scheme analyzed in this article is an intriguing option for further improving the spectral efficiency of these types of networks. Furthermore, to properly evaluate the full-duplex self-backhauling solution for the AN, its performance is compared to two reference schemes, one of which relies on traditional half-duplex communication while in the other the AN acts as a one-directional full-duplex relay. The latter reference scheme corresponds to the solution mostly investigated in the earlier works.

In addition, in this article it is also assumed that the AN has large arrays of antennas at its disposal. Therefore, in the full-duplex solution, the same time-frequency resource can be used for all the individual UL and DL transmissions, as well as for the wireless backhaul link, since such massive antenna arrays allow for efficient beamforming, which can be used to prevent the interference between the various spatial streams [24], [30], [33]. The massive arrays also facilitate efficient attenuation of the SI by zero-forcing (ZF) beamforming [9], [38]. Namely, the transmit signals will be precoded such that nulls are formed in the positions of all the receive antennas, which will significantly decrease the SI power coupled back to the receivers. To suppress the residual SI remaining after the ZF procedure, additional SI cancellation can also be performed, e.g., in the digital domain [3], [6], [8], [17].

The different communication schemes are then analyzed under a scenario where a minimum Quality-of-Service (QoS) requirement is given for each UE, defined in terms of minimum DL and UL data rates. This definition ensures uniform QoS for all the UEs, which makes it a reasonable choice. The problem is then to determine the minimum transmit powers for each communicating party under the constraint that each UE achieves the minimum required data rate. Furthermore, since wireless self-backhauling is assumed, the AN and the BN must also allocate some transmit power for the backhaul link to ensure sufficient backhauling capability. A similar system was considered by the authors already in [33], where the sum-rate was optimized under a greatly simplified system model, while the transmit power minimization under QoS constraints was preliminarily considered in [24]. The current article completes and archives our research work in the most comprehensive form under a generic setting by presenting closed-form solutions for the optimal transmit powers in three different communication schemes: a full-duplex scheme, a half-duplex scheme, and a hybrid relay scheme. To the best of our knowledge, this is something that has not been solved before for any self-backhauling radio access system.

The major contributions of this paper can be detailed as follows:

- We derive closed-form solutions for the optimal transmit powers of all the considered communication schemes that fulfill the QoS requirements.
- We show that the full-duplex and hybrid relay schemes cannot always fulfill the QoS requirements, even if the transmit powers tend towards infinity. In other words, these two schemes are feasible only under some circumstances, meaning that there is a fundamental limit for the data rates that they can achieve. The condition for this is derived

in closed form, while accurate approximations for the feasibility boundary are also provided.
- We provide extensive numerical results to illustrate different aspects of the considered communication schemes. In particular, the numerical results show that in most cases the full-duplex scheme is indeed the most transmit power efficient solution. However, the results also indicate that the schemes utilizing a full-duplex capable AN are fundamentally limited to relatively small cell sizes.

The rest of this article is organized as follows. The system model is first presented in Section II, alongside with the achievable DL and UL data rate expressions of the three different communication schemes. The optimal QoS fulfilling transmit powers are then derived in Section III. After this, the feasibility of the full-duplex and hybrid relay schemes is investigated in Section IV, the feasibility boundaries being derived in closed form. The numerical results are then given and analyzed in Section V, while the conclusions are drawn in Section VI.

II. System Model, Communication Schemes, and Sum-Rate Expressions

Let us consider a wireless cell with a large-array AN that is communicating with a multiple-input and multiple-output (MIMO) BN and half-duplex single-antenna UEs, the UEs being further divided into UL and DL UEs. The AN is assumed to have N_t transmit and N_r receive antennas, while the amount of transmitted and received signal streams is assumed to be significantly lower. Moreover, the same antenna arrays are used for serving the DL and UL UEs, as well as for communicating with the BN. Three different communication schemes are analyzed in this article, each of them depicted in Fig. 1. Below, we describe the different communication schemes in detail, and also derive the expressions of the achievable data rates for each scheme. These can then be used to derive the optimal transmit power allocations.

A. Full-Duplex Scheme

In the full-duplex scheme, the AN transmits signals simultaneously to the BN and to the DL UEs while also receiving signals from the UL UEs and the BN, all of the transmissions occurring on the same center frequency. Consequently, both the AN and the BN must be full-duplex capable, while the UEs are legacy half-duplex devices, as already mentioned. This type of a full-duplex system suffers from the SI, the IUI between the UL and the DL UEs, as well as from the interference between the BN and the UEs. Even though there are also advanced methods for attenuating the UL-to-DL IUI [39]–[41], in this work we assume that its power level is only affected by the transmit power of the UL UEs and the path losses between the UL and DL UEs.

Denoting the number of DL UEs by D and the number of transmitted backhaul signal streams by M_t^B, the overall stacked spatial signal received by the UEs and the BN can be represented as a vector, whose first D elements contain the samples received by the DL UEs, while the last M_t^B elements contain the samples received by the BN (consisting of the parallel streams of backhauled UL data). This total received signal vector can be written as follows:

$$\mathbf{y} = \mathbf{L}_\mathrm{t}\mathbf{H}_\mathrm{t}\mathbf{x} + \mathbf{z}, \tag{1}$$

where $\mathbf{L}_\mathrm{t} = \mathrm{diag}\left(\sqrt{L_1^\mathrm{d}}, \ldots, \sqrt{L_D^\mathrm{d}}, \ldots, \sqrt{L_\mathrm{B}}, \ldots, \sqrt{L_\mathrm{B}}\right)$ is a $(D + M_\mathrm{t}^\mathrm{B}) \times (D + M_\mathrm{t}^\mathrm{B})$ diagonal matrix, L_i^d is the path loss normalized fading variance between the AN and the ith DL UE, L_B is the path loss normalized fading variance between the AN and the BN, $\mathbf{H}_\mathrm{t} \in \mathbb{C}^{(D+M_\mathrm{t}^\mathrm{B})\times N_\mathrm{t}}$ is the normalized channel matrix between the AN and all the intended receivers, $\mathbf{x} \in \mathbb{C}^{N_\mathrm{t}\times 1}$ is the transmit signal of the AN and $\mathbf{z} \in \mathbb{C}^{(D+M_\mathrm{t}^\mathrm{B})\times 1}$ represents the different noise and interference sources. In this article, Rayleigh fading between all communicating parties is assumed, which means that the entries of $\mathbf{H}_\mathrm{t}$ are independent and identically distributed (i.i.d.) zero-mean complex Gaussian random variables with unit variance. In the continuation, to simplify the literary presentation, the path loss normalized fading variances are simply referred to as path losses. Also note that, while the path losses between the AN and the UEs are different, the path losses of the backhaul signals are the same as they all correspond to the link between the AN and the BN.

The precoded transmit signal $\mathbf{x}$ is formed from the DL and backhaul transmit data as follows:

$$\mathbf{x} = \mathbf{W}\mathbf{\Gamma}\mathbf{q}, \tag{2}$$

where $\mathbf{W} \in \mathbb{C}^{N_\mathrm{t}\times(D+M_\mathrm{t}^\mathrm{B})}$ is the precoding matrix, $\mathbf{q} \in \mathbb{C}^{(D+M_\mathrm{t}^\mathrm{B})\times 1}$ contains all the transmit data symbols, $\mathbf{\Gamma} = \mathrm{diag}\left(\sqrt{p_1^\mathrm{d}}, \ldots, \sqrt{p_D^\mathrm{d}}, \ldots, \sqrt{P_\mathrm{u}^\mathrm{B}/M_\mathrm{t}^\mathrm{B}}, \ldots, \sqrt{P_\mathrm{u}^\mathrm{B}/M_\mathrm{t}^\mathrm{B}}\right)$ is a $(D+M_\mathrm{t}^\mathrm{B}) \times (D + M_\mathrm{t}^\mathrm{B})$ diagonal matrix, p_i^d is the transmit power allocated from the ith DL signal stream, and P_u^B is the total transmit power allocated for backhauling the UL data. The power of the data symbols is assumed to be normalized as $\mathrm{E}\left[|q_i|^2\right] = 1$ where q_i is the ith element of $\mathbf{q}$. Even though the transmitter's power amplifier–induced nonlinear distortion is typically a significant issue in full-duplex devices [3], in this analysis we are using a linear signal model for simplicity. In fact, in a massive MIMO transmitter, the powers of the individual transmitters are typically small, somewhat alleviating the effects of the nonlinearities.

The precoding is performed using ZF beamforming since it typically performs well under high signal-to-noise ratios [42]. Assuming that also the effective SI channel experiences Rayleigh fading, the SI channel matrix between the AN transmitters and receivers can be expressed as $\mathbf{L}_\mathrm{s}\mathbf{H}_\mathrm{s} \in \mathbb{C}^{N_\mathrm{r}\times N_\mathrm{t}}$, where $\mathbf{L}_\mathrm{s}$ is a diagonal matrix containing the amounts of SI suppression between the transmitter and receiver pairs without any ZF nulling (assuming that the amounts of SI suppression are equal for all transmitter and receiver pairs) and the elements of $\mathbf{H}_\mathrm{s}$ are i.i.d. zero-mean complex Gaussian random variables with unit variance. Note that assuming the SI channel to experience Rayleigh fading can be expected to be relatively accurate when there is a certain level of active SI suppression before the total received signal is decoded [1]. For the same reason, the overall SI suppression between each transmitter and receiver pair can be expected to be roughly the same, as more

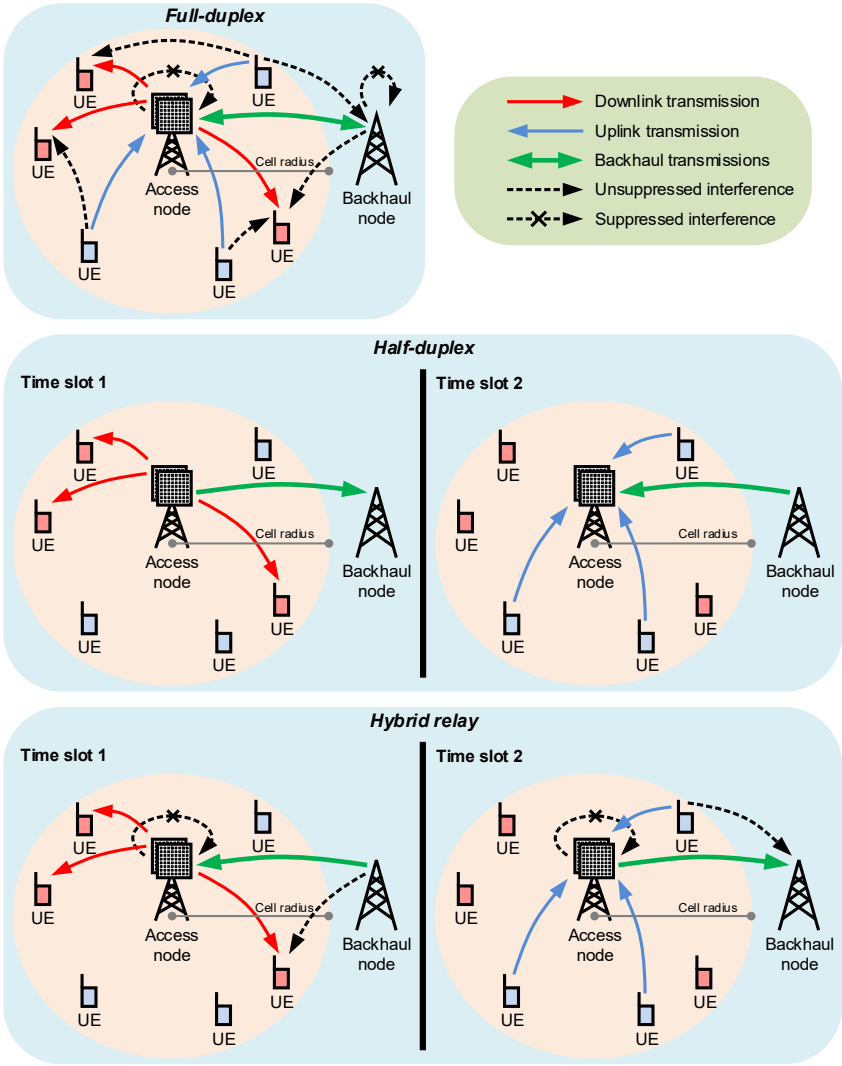


Fig. 1: An illustration of the three considered communication schemes: the *full-duplex* scheme, the *half-duplex* scheme, and the *hybrid relay* scheme.

active SI cancellation is typically obtained in the receivers with stronger SI coupling in the antenna interface [1].

Under the condition that the AN has full knowledge of the channel state information (CSI) of the links used for data transfer (that is, the channel matrix $\mathbf{H}_{\mathrm{t}}$ is fully known) and assuming that $N_t > N_r+D+M_{\mathrm{t}}^{\mathrm{B}}$, the transmitter ZF precoding matrix in full-duplex mode can then be written as [42]

$$\mathbf{W} = \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} \mathbf{\Lambda}, \tag{3}$$

where $\mathbf{H}^H = \begin{bmatrix} \mathbf{H}_{\mathrm{t}}^H & \widehat{\mathbf{H}}_{\mathrm{s}}^H \end{bmatrix}$, $\widehat{\mathbf{H}}_{\mathrm{s}}$ is an imperfect estimate of the effective SI coupling channel, and $\mathbf{\Lambda} \in \mathbb{C}^{(D+M_{\mathrm{t}}^{\mathrm{B}}+N_{\mathrm{r}})\times(D+M_{\mathrm{t}}^{\mathrm{B}})}$ is a non-square diagonal normalization matrix containing the normalization factor $\sqrt{N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}} - N_{\mathrm{r}}}$ in each diagonal [38], [42]. Moreover, $(\cdot)^H$ denotes the Hermitian transpose. The purpose of the normalization matrix is simply to ensure that precoding does not affect the expected effective powers of the data symbols.

It should be noted that assuming the AN to have full knowledge of the CSI of the data transfer links is obviously not fully practical, but it allows the derivation of analytical data rate expressions that provide information about the ultimate performance limits of the considered system. Namely, this assumption means that, apart from SI, none of the signals received or transmitted by the AN interfere with each other, representing a best-case scenario in this respect. Nevertheless, the effect of residual SI is still considered in the system, as no full knowledge of the effective SI coupling channel is assumed. Furthermore, in order to simplify the system models

and derivations, only the beamforming performed by the AN is explicitly considered, meaning that analysis of any potential spatial domain processing in the other nodes is omitted.

Now, we can rewrite the signals received by the DL UEs and the BN as

$$\mathbf{y} = \mathbf{L}_\mathrm{t}\mathbf{H}_\mathrm{t}\mathbf{x} + \mathbf{z} = \mathbf{L}_\mathrm{t}\mathbf{H}_\mathrm{t}\mathbf{W}\boldsymbol{\Gamma}\mathbf{q} + \mathbf{z} = \mathbf{L}_\mathrm{t}\widetilde{\boldsymbol{\Lambda}}\boldsymbol{\Gamma}\mathbf{q} + \mathbf{z}, \quad (4)$$

where $\widetilde{\boldsymbol{\Lambda}} \in \mathbb{C}^{(D+M_\mathrm{t}^\mathrm{B})\times(D+M_\mathrm{t}^\mathrm{B})}$ denotes $\boldsymbol{\Lambda}$ with all the zero rows removed. Note that the noise-plus-interference signal $\mathbf{z}$ includes here also the residual SI resulting from the imperfect nulling of the receive antennas by the precoder. To express the individual received data streams, (4) can alternatively be written component wise using the elements of the various matrices. Then, we get

$$y_i = \begin{cases} \sqrt{L_i^\mathrm{d}\left(N_\mathrm{t} - D - M_\mathrm{t}^\mathrm{B} - N_\mathrm{r}\right)p_i^\mathrm{d}}q_i + z_i, \; i = 1, \ldots, D \\ \sqrt{L_\mathrm{B}(N_\mathrm{t} - D - M_\mathrm{t}^\mathrm{B} - N_\mathrm{r})\frac{P_\mathrm{u}^\mathrm{B}}{M_\mathrm{t}^\mathrm{B}}}q_i + z_i, \; i = D+1, \ldots, D+M_\mathrm{t}^\mathrm{B} \end{cases} \quad (5)$$

Stemming from (5) and assuming a large transmit antenna array [38], [42], the signal-to-interference-plus-noise ratio (SINR) of the ith DL signal can then be expressed as follows for the full-duplex scheme:

$$\begin{aligned} \mathit{SINR}_i^{\mathrm{d,FD}} &= \frac{\mathrm{E}\left[\left|y_i - z_i\right|^2\right]}{\mathrm{E}\left[\left|z_i\right|^2\right]} \\ &= \frac{\mathrm{E}\left[\left|\sqrt{L_i^\mathrm{d}\left(N_\mathrm{t} - D - M_\mathrm{t}^\mathrm{B} - N_\mathrm{r}\right)p_i^\mathrm{d}}q_i\right|^2\right]}{\mathrm{E}\left[\left|z_i\right|^2\right]} \\ &= \frac{\Lambda_\mathrm{t}^\mathrm{FD}L_i^\mathrm{d}p_i^\mathrm{d}}{\sigma_\mathrm{n}^2 + L_i^\mathrm{Bd}P_\mathrm{d}^\mathrm{B} + \sum_{j=1}^{U}L_{ij}^\mathrm{ud}p_j^\mathrm{u}}, \quad i = 1, \ldots, D, \end{aligned} \quad (6)$$

where the power of the noise-plus-interference term z_i has been expanded to reflect the various components, $\Lambda_\mathrm{t}^\mathrm{FD} = N_\mathrm{t} - N_\mathrm{r} - D - M_\mathrm{t}^\mathrm{B}$, and the rest of the symbols are as defined in Table I. To further illustrate the many symbols and parameters used throughout this work, Fig. 2 provides also a visual depiction of their meaning within the considered system.

Similarly, the SINR of the backhaul signal streams transmitted by the AN, used for backhauling the UL data, can be written as follows

$$\mathit{SINR}_\mathrm{B}^{\mathrm{u,FD}} = \frac{\Lambda_\mathrm{t}^\mathrm{FD}L_\mathrm{B}P_\mathrm{u}^\mathrm{B}}{M_\mathrm{t}^\mathrm{B}\left(\sigma_\mathrm{n}^2 + \alpha_\mathrm{BN}P_\mathrm{d}^\mathrm{B} + \sum_{j=1}^{U}L_j^\mathrm{Bu}p_j^\mathrm{u}\right)}, \quad (7)$$

with the symbols again defined in Table I and illustrated in Fig. 2.

The SINRs of the signals *received* by the AN can be derived in an essentially similar manner as those of the transmit signals (cf. [38]), and hence their detailed derivation is omitted for brevity. In particular, the SINR of the jth UL signal can be shown to read:

$$\mathit{SINR}_j^{\mathrm{u,FD}} = \frac{\Lambda_\mathrm{r}^\mathrm{FD}L_j^\mathrm{u}p_j^\mathrm{u}}{\sigma_\mathrm{n}^2 + \alpha_\mathrm{AN}\left(P_\mathrm{u}^\mathrm{B} + \sum_{i=1}^{D}p_i^\mathrm{d}\right)}, \quad j = 1, \ldots, U, \quad (8)$$

TABLE I: The most important symbols used in the paper.

Variable	Definition
N_t / N_r	Number of transmit/receive antennas at the AN
M_t^B / M_r^B	Number of backhaul signal streams transmitted/received by the AN
D / U	Number of DL/UL UEs in the cell
$\Lambda_\mathrm{t}^\mathrm{X}$ / $\Lambda_\mathrm{r}^\mathrm{X}$	The degrees-of-freedom of the AN transmitter/receiver, $\mathrm{X} = \{\mathrm{FD}, \mathrm{HD}, \mathrm{RL}\}$
L_i^d / L_j^u	Path loss between the AN and the ith DL / jth UL UE
L_{ij}^ud	Path loss between the ith DL and the jth UL UE
L_i^Bd / L_j^Bu	Path loss between the BN and the ith DL / jth UL UE
L_B	Path loss between the AN and the BN
σ_n^2	Noise floor in all the receivers
$\alpha_\mathrm{AN}/\alpha_\mathrm{BN}$	Amount of SI cancellation in the AN/BN
p_i^d	Transmit power used for the ith DL signal stream
p_j^u	Transmit power of the jth UL UE
P_d^B	Total amount of transmit power used by the BN
P_u^B	Total amount of transmit power allocated for self-backhauling in the AN
η	Proportion of time spent in the DL time slot (in the half-duplex and hybrid relay schemes)
ρ_d / ρ_u	DL/UL rate requirement of an individual UE

where again the symbols are as defined in Table I, and $\Lambda_\mathrm{r}^\mathrm{FD} = N_\mathrm{r} - U - M_\mathrm{r}^\mathrm{B}$. Correspondingly, the SINR of the backhaul signals received by the AN, backhauling the DL data, is as follows:

$$\mathit{SINR}_\mathrm{B}^{\mathrm{d,FD}} = \frac{\Lambda_\mathrm{r}^\mathrm{FD}L_\mathrm{B}P_\mathrm{d}^\mathrm{B}}{M_\mathrm{r}^\mathrm{B}\left[\sigma_\mathrm{n}^2 + \alpha_\mathrm{AN}\left(P_\mathrm{u}^\mathrm{B} + \sum_{i=1}^{D}p_i^\mathrm{d}\right)\right]}. \quad (9)$$

The hereby obtained SINR expressions can then be used to determine the achievable rates of the full-duplex scheme. In particular, using (6), the DL sum-rate of this communication scheme can be expressed as follows:

$$\begin{aligned} R^\mathrm{d} &= \sum_{i=1}^{D}R_i^\mathrm{d}, \\ R_i^\mathrm{d} &= \log_2\left(1 + \mathit{SINR}_i^{\mathrm{d,FD}}\right) \\ &= \log_2\left(1 + \frac{\Lambda_\mathrm{t}^\mathrm{FD}L_i^\mathrm{d}p_i^\mathrm{d}}{\sigma_\mathrm{n}^2 + L_i^\mathrm{Bd}P_\mathrm{d}^\mathrm{B} + \sum_{j=1}^{U}L_{ij}^\mathrm{ud}p_j^\mathrm{u}}\right), \end{aligned} \quad (10)$$

It can be observed that, in this communication scheme, the DL data rate is degraded by the UL-to-DL IUI and by the interference produced by the BN transmission. Similarly, using (8), the total UL data rate can be written as:

$$\begin{aligned} R^\mathrm{u} &= \sum_{j=1}^{U}R_j^\mathrm{u}, \\ R_j^\mathrm{u} &= \log_2\left(1 + \mathit{SINR}_j^{\mathrm{u,FD}}\right) \\ &= \log_2\left(1 + \frac{\Lambda_\mathrm{r}^\mathrm{FD}L_j^\mathrm{u}p_j^\mathrm{u}}{\sigma_\mathrm{n}^2 + \alpha_\mathrm{AN}\left(P_\mathrm{u}^\mathrm{B} + \sum_{i=1}^{D}p_i^\mathrm{d}\right)}\right), \end{aligned} \quad (11)$$

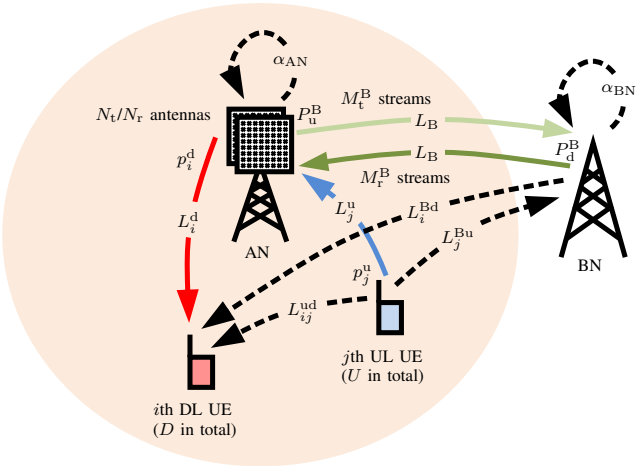


Fig. 2: An illustration depicting the relevant symbols.

where the SINR is now degraded by the residual SI within the AN. Note that this work does not assume any specific SI cancellation performance since all the derivations are done for an arbitrary amount of SI cancellation, consisting of passive antenna isolation, ZF beamforming at the transmit side to form nulls at the receive antennas, and other possible SI suppression methods.

Since the AN must also be capable of backhauling all the data, the backhaul data rates must also be taken into account in the analysis. With the help of (7), the data rate of the backhaul signal transmitted by the AN (for backhauling UL data) can be expressed as follows:

$$\begin{aligned} R^{\mathrm{u,B}} &= \sum_{k=1}^{M_{\mathrm{t}}^{\mathrm{B}}} \log_2\left(1+\mathit{SINR}_{\mathrm{B}}^{\mathrm{u,FD}}\right) \\ &= M_{\mathrm{t}}^{\mathrm{B}} \log_2\left(1+\frac{\Lambda_{\mathrm{t}}^{\mathrm{FD}} L_{\mathrm{B}} P_{\mathrm{u}}^{\mathrm{B}}/M_{\mathrm{t}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2+\alpha_{\mathrm{BN}} P_{\mathrm{d}}^{\mathrm{B}}+\sum_{j=1}^{U} L_j^{\mathrm{Bu}} p_j^{\mathrm{u}}}\right), \end{aligned} \tag{12}$$

which is affected by the residual SI within the BN, and also by the interference caused by the UL transmissions. In a similar fashion, using (9), the data rate of the received backhaul signal streams in the AN (for backhauling DL data) can be written as follows:

$$\begin{aligned} R^{\mathrm{d,B}} &= \sum_{l=1}^{M_{\mathrm{r}}^{\mathrm{B}}} \log_2\left(1+\mathit{SINR}_{\mathrm{B}}^{\mathrm{d,FD}}\right) \\ &= M_{\mathrm{r}}^{\mathrm{B}} \log_2\left(1+\frac{\Lambda_{\mathrm{r}}^{\mathrm{FD}} L_{\mathrm{B}} P_{\mathrm{d}}^{\mathrm{B}}/M_{\mathrm{r}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2+\alpha_{\mathrm{AN}}\left(P_{\mathrm{u}}^{\mathrm{B}}+\sum_{i=1}^{D} p_i^{\mathrm{d}}\right)}\right). \end{aligned} \tag{13}$$

The data rate of the received backhaul signals is decreased by the residual SI within the AN, similar to the UL signals. Put together, the data rate expressions in (10)–(13) can be used to determine the optimal transmit powers for the considered system under some given data rate requirements, as will be done in Section III.

B. Half-duplex Scheme

Perhaps the most obvious alternative to the aforementioned full-duplex scheme is for all the nodes to operate in half-duplex mode. This can be done by utilizing time-division duplexing (TDD) where each node within the considered system either transmits or receives at any given time, using all of the available spectrum. In terms of the analyzed AN with a wireless backhaul link, one possible TDD scheme is shown in the middle part of Fig. 1. There, the system has two different time slots: one where the AN transmits data to the BN and to the UEs, and one where it receives data from them. As can easily be observed from Fig. 1, this type of a scheme requires only half-duplex capable nodes since none of them have to engage in simultaneous transmission and reception. This removes the problems of SI and UL-to-DL IUI at the cost of decreased spectral efficiency since the AN now requires more temporal resources to carry out the same tasks in comparison to the full-duplex scheme.

Although the detailed derivations must be omitted for brevity, the SINRs for this type of a half-duplex scheme can be obtained in a similar manner as done for the full-duplex scheme above, and the corresponding DL sum-rate of the half-duplex system can be expressed as follows:

$$\begin{aligned} R^{\mathrm{d}} &= \sum_{i=1}^{D} R_i^{\mathrm{d}}, \\ R_i^{\mathrm{d}} &= \eta \log_2\left(1+\mathit{SINR}_i^{\mathrm{d,HD}}\right) \\ &= \eta \log_2\left(1+\frac{\Lambda_{\mathrm{t}}^{\mathrm{HD}} L_i^{\mathrm{d}} p_i^{\mathrm{d}}}{\sigma_{\mathrm{n}}^2}\right), \end{aligned} \tag{14}$$

where the symbols are as defined in Table I and $\Lambda_{\mathrm{t}}^{\mathrm{HD}} = N_{\mathrm{t}} - D - M_{\mathrm{t}}^{\mathrm{B}}$. The relative lengths of the two time slots are controlled by the duplexing parameter η, which defines the proportion of time spent in the DL time slot (the relative length of the UL time slot being $1-\eta$). The corresponding

UL sum-rate can then be written as:

$$\begin{aligned} R^{\mathrm{u}} &= \sum_{j=1}^{U} R_j^{\mathrm{u}}, \\ R_j^{\mathrm{u}} &= (1-\eta)\log_2\left(1+\mathit{SINR}_j^{\mathrm{u,HD}}\right) \\ &= (1-\eta)\log_2\left(1+\frac{\Lambda_{\mathrm{r}}^{\mathrm{HD}} L_j^{\mathrm{u}} p_j^{\mathrm{u}}}{\sigma_{\mathrm{n}}^2}\right), \end{aligned} \tag{15}$$

where $\Lambda_{\mathrm{r}}^{\mathrm{HD}} = N_{\mathrm{r}} - U - M_{\mathrm{r}}^{\mathrm{B}}$. Hence, as can be observed, the data rates are decreased due to time division, but the DL and UL transmissions in the half-duplex scheme do not suffer from any form of interference. Furthermore, no degrees-of-freedom are lost due to having to null the receive antennas. The backhaul data rates in half-duplex mode can then be expressed as

$$\begin{aligned} R^{\mathrm{u,B}} &= \sum_{k=1}^{M_{\mathrm{t}}^{\mathrm{B}}} \eta\log_2\left(1+\mathit{SINR}_{\mathrm{B}}^{\mathrm{u,HD}}\right) \\ &= \eta M_{\mathrm{t}}^{\mathrm{B}}\log_2\left(1+\frac{\Lambda_{\mathrm{t}}^{\mathrm{HD}} L_{\mathrm{B}} P_{\mathrm{u}}^{\mathrm{B}}}{M_{\mathrm{t}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}\right), \end{aligned} \tag{16}$$

$$\begin{aligned} R^{\mathrm{d,B}} &= \sum_{l=1}^{M_{\mathrm{r}}^{\mathrm{B}}} (1-\eta)\log_2\left(1+\mathit{SINR}_{\mathrm{B}}^{\mathrm{d,HD}}\right) \\ &= (1-\eta)\, M_{\mathrm{r}}^{\mathrm{B}}\log_2\left(1+\frac{\Lambda_{\mathrm{r}}^{\mathrm{HD}} L_{\mathrm{B}} P_{\mathrm{d}}^{\mathrm{B}}}{M_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}\right). \end{aligned} \tag{17}$$

Based on (14)–(17), an intuitive interpretation regarding the relationship between the data rates of the full-duplex and half-duplex schemes is that it highly depends on the level of the total interference. With low path loss between the UL and DL UEs and/or poor SI cancellation performance, the half-duplex scheme is likely to be the better option due to it being immune to the interference. On the other hand, if the AN is capable of efficiently suppressing the SI signal and the UEs do not strongly interfere with each other or with the BN, the full-duplex scheme will likely provide the higher performance. These aspects are investigated further in Section V with the help of numerical results.

C. Hybrid Relay Scheme

In addition to the above extreme cases of purely full-duplex and half-duplex systems, an interesting scheme is a full-duplex relay-type AN, which simply relays the UL signal to the BN during one time slot, and then in the other time slot relays the signal from the BN to the DL UEs. The bottom part of Fig. 1 illustrates this type of a solution. The benefit of this scheme is that the problem of UL-to-DL IUI is completely avoided, similar to the half-duplex scheme, while the full-duplex capability of the AN is still utilized to some extent as the relaying is performed inband. The obvious drawback is, however, that now everything cannot be done simultaneously, which will inherently decrease the achievable rate. Furthermore, the relay scheme still suffers from the interference between the BN and the UEs.

Also the SINRs of this type of a hybrid relay scheme can be derived in a similar manner as those of the full-duplex scheme in Section II-A, the DL sum-rate being now

$$\begin{aligned} R^{\mathrm{d}} &= \sum_{i=1}^{D} R_i^{\mathrm{d}}, \\ R_i^{\mathrm{d}} &= \eta\log_2\left(1+\mathit{SINR}_i^{\mathrm{d,RL}}\right) \\ &= \eta\log_2\left(1+\frac{\Lambda_{\mathrm{t}}^{\mathrm{RL}} L_i^{\mathrm{d}} p_i^{\mathrm{d}}}{\sigma_{\mathrm{n}}^2 + L_i^{\mathrm{Bd}} P_{\mathrm{d}}^{\mathrm{B}}}\right), \end{aligned} \tag{18}$$

where $\Lambda_{\mathrm{t}}^{\mathrm{RL}} = N_{\mathrm{t}} - D - N_{\mathrm{r}}$. The UL sum-rate can correspondingly be written as:

$$\begin{aligned} R^{\mathrm{u}} &= \sum_{j=1}^{U} R_j^{\mathrm{u}}, \\ R_j^{\mathrm{u}} &= (1-\eta)\log_2\left(1+\mathit{SINR}_j^{\mathrm{u,RL}}\right) \\ &= (1-\eta)\log_2\left(1+\frac{\Lambda_{\mathrm{r}}^{\mathrm{RL}} L_j^{\mathrm{u}} p_j^{\mathrm{u}}}{\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{AN}} P_{\mathrm{u}}^{\mathrm{B}}}\right), \end{aligned} \tag{19}$$

where $\Lambda_{\mathrm{r}}^{\mathrm{RL}} = N_{\mathrm{r}} - U$. Now, there is still some interference, which degrades the DL and UL SINRs, but this scheme can be considered a trade-off between tolerating interference and duplexing the transmissions and receptions in time.

Finally, the backhaul data rates of the hybrid relay scheme can be expressed as follows:

$$\begin{aligned} R^{\mathrm{u,B}} &= \sum_{k=1}^{M_{\mathrm{t}}^{\mathrm{B}}} (1-\eta)\log_2\left(1+\mathit{SINR}_{\mathrm{B}}^{\mathrm{u,RL}}\right) \\ &= (1-\eta)\, M_{\mathrm{t}}^{\mathrm{B}}\log_2\left(1+\frac{\Lambda_{\mathrm{t,B}}^{\mathrm{RL}} L_{\mathrm{B}} P_{\mathrm{u}}^{\mathrm{B}}/M_{\mathrm{t}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2 + \sum_{j=1}^{U} L_j^{\mathrm{Bu}} p_j^{\mathrm{u}}}\right), \end{aligned} \tag{20}$$

$$\begin{aligned} R^{\mathrm{d,B}} &= \sum_{l=1}^{M_{\mathrm{r}}^{\mathrm{B}}} \eta\log_2\left(1+\mathit{SINR}_{\mathrm{B}}^{\mathrm{d,RL}}\right) \\ &= \eta M_{\mathrm{r}}^{\mathrm{B}}\log_2\left(1+\frac{\Lambda_{\mathrm{r,B}}^{\mathrm{RL}} L_{\mathrm{B}} P_{\mathrm{d}}^{\mathrm{B}}/M_{\mathrm{r}}^{\mathrm{B}}}{\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{AN}} \sum_{i=1}^{D} p_i^{\mathrm{d}}}\right), \end{aligned} \tag{21}$$

where $\Lambda_{\mathrm{t,B}}^{\mathrm{RL}} = N_{\mathrm{t}} - N_{\mathrm{r}} - M_{\mathrm{t}}^{\mathrm{B}}$ and $\Lambda_{\mathrm{r,B}}^{\mathrm{RL}} = N_{\mathrm{r}} - M_{\mathrm{r}}^{\mathrm{B}}$ are the degrees-of-freedom of the AN transmitter and receiver for backhauling data in the hybrid relay scheme, respectively. Again, a crucial aspect of the considered cell is that the backhaul link should be able to match the data rates of UL and DL. Otherwise the system will be bottlenecked by the backhaul connection.

III. Transmit Sum-Power Minimization under Rate Constraints

Next, the problem of minimizing the transmit powers of the system under some given data rate requirements for the DL and the UL is investigated. In particular, let us define the per-UE QoS requirements in terms of the minimum data rates as ρ_{d} and ρ_{u} for the DL and the UL, respectively. This results in the following optimization problem.

Problem (Transmit Sum-Power Minimization):

$$\begin{aligned} \underset{\mathbf{p},\, P_{\mathrm{d}}^{\mathrm{B}},\, P_{\mathrm{u}}^{\mathrm{B}}}{\text{minimize}} \quad & \left(\mathbf{1}_{D+U}^T\mathbf{p} + P_{\mathrm{d}}^{\mathrm{B}} + P_{\mathrm{u}}^{\mathrm{B}}\right) \\ \text{subject to} \quad & \text{C1: } R_i^{\mathrm{d}} \geq \rho_{\mathrm{d}},\ i = 1,\ldots,D, \\ & \text{C2: } R_j^{\mathrm{u}} \geq \rho_{\mathrm{u}},\ j = 1,\ldots,U, \\ & \text{C3: } R^{\mathrm{d,B}} \geq \sum_{i=1}^{D} R_i^{\mathrm{d}}, \\ & \text{C4: } R^{\mathrm{u,B}} \geq \sum_{j=1}^{U} R_j^{\mathrm{u}}, \end{aligned} \tag{22}$$

where $\mathbf{p} = \begin{bmatrix}\mathbf{p}_{\mathrm{d}}^T & \mathbf{p}_{\mathrm{u}}^T\end{bmatrix}^T$, $\mathbf{p}_{\mathrm{d}}$ and $\mathbf{p}_{\mathrm{u}}$ are column vectors containing the DL and UL transmit powers p_i^{d} and p_j^{u} stacked, respectively, and $\mathbf{1}_N$ is a column vector consisting of N ones. The constraints C1 and C2 ensure the QoS of the UEs, while the constraints C3 and C4 ensure sufficient backhauling capability in the AN. This optimization problem will next be solved separately for the three considered communication schemes, while the associated infeasible system scenarios and QoS requirements that manifest themselves as negative transmit power values in the following theorems are characterized later in Section IV.

A. Full-Duplex Scheme

Theorem 1: The optimal DL and UL transmit powers for the full-duplex scheme are

$$\begin{aligned} \mathbf{p}^* &= \begin{bmatrix}\mathbf{p}_{\mathrm{d}}^* \\ \mathbf{p}_{\mathrm{u}}^*\end{bmatrix} \\ &= \begin{bmatrix} \frac{\sigma_{\mathrm{u}}^2\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}\mathbf{q}_{\mathrm{d}} + \frac{\sigma_{\mathrm{n}}^2\left(\frac{1+S_{\mathrm{d}}\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}+\gamma_{\mathrm{t}}^{\mathrm{B}}\right)}{\theta(1-\gamma^{\mathrm{B}})}\left(\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}\mathbf{q}_{\mathrm{B/d}} + \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}}}{\alpha_{\mathrm{AN}}}\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}\mathbf{q}_{\mathrm{u}}\right) \\ \frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{u}}\left(\frac{1+S_{\mathrm{d}}\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}+\gamma_{\mathrm{t}}^{\mathrm{B}}\right)}{\theta(1-\gamma^{\mathrm{B}})}\mathbf{q}_{\mathrm{u}} \end{bmatrix}, \end{aligned} \tag{23}$$

when each element of $\mathbf{p}^*$ is positive and finite. Otherwise the QoS requirements cannot be fulfilled and the system is infeasible. Here, $\gamma_{\mathrm{d}} = \frac{\alpha_{\mathrm{AN}}(2^{\rho_{\mathrm{d}}}-1)}{\Lambda_{\mathrm{t}}^{\mathrm{FD}}}$, $\gamma_{\mathrm{u}} = \frac{\alpha_{\mathrm{AN}}(2^{\rho_{\mathrm{u}}}-1)}{\Lambda_{\mathrm{r}}^{\mathrm{FD}}}$, $\gamma_{\mathrm{r}}^{\mathrm{B}} = \frac{M_{\mathrm{r}}^{\mathrm{B}}\left(2^{D\rho_{\mathrm{d}}/M_{\mathrm{r}}^{\mathrm{B}}}-1\right)}{\Lambda_{\mathrm{r}}^{\mathrm{FD}}L_{\mathrm{B}}}$, $\gamma_{\mathrm{t}}^{\mathrm{B}} = \frac{M_{\mathrm{t}}^{\mathrm{B}}\left(2^{U\rho_{\mathrm{u}}/M_{\mathrm{t}}^{\mathrm{B}}}-1\right)}{\Lambda_{\mathrm{t}}^{\mathrm{FD}}L_{\mathrm{B}}}$, $\gamma^{\mathrm{B}} = \alpha_{\mathrm{AN}}\alpha_{\mathrm{BN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}$, $\{\mathbf{q}_{\mathrm{d}}\}_i = 1/L_i^{\mathrm{d}}$, $\{\mathbf{q}_{\mathrm{u}}\}_j = 1/L_j^{\mathrm{u}}$, $\{\mathbf{q}_{\mathrm{B/d}}\}_i = L_i^{\mathrm{Bd}}/L_i^{\mathrm{d}}$, $\{\mathbf{L}_{\mathrm{Bu}}\}_j = L_j^{\mathrm{Bu}}$, $\{\mathbf{L}_{\mathrm{ud}}\}_{ij} = L_{ij}^{\mathrm{ud}}$, $\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}} = \mathbf{L}_{\mathrm{ud}} \circ \mathbf{q}_{\mathrm{d}}\mathbf{1}_{\mathbf{U}}^T$, and $\circ$ denotes the Hadamard product. Note that $\mathbf{q}_{\mathrm{d}}$, $\mathbf{q}_{\mathrm{u}}$, $\mathbf{q}_{\mathrm{B/d}}$, and $\mathbf{L}_{\mathrm{Bu}}$ are column vectors. Furthermore, the parameter θ is defined as

$$\theta = 1 - \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{1-\gamma^{\mathrm{B}}} - \frac{\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}{1-\gamma^{\mathrm{B}}} - \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}}S_{\mathrm{ud}}}{\alpha_{\mathrm{AN}}\left(1-\gamma^{\mathrm{B}}\right)}, \tag{24}$$

where $S_{\mathrm{d}} = \mathbf{1}_D^T\mathbf{q}_{\mathrm{d}}$, $S_{\mathrm{B/d}} = \mathbf{1}_D^T\mathbf{q}_{\mathrm{B/d}}$, $S_{\mathrm{B/u}} = \mathbf{L}_{\mathrm{Bu}}^T\mathbf{q}_{\mathrm{u}}$, and $S_{\mathrm{ud}} = \mathbf{1}_D^T\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}\mathbf{q}_{\mathrm{u}}$. The optimal backhauling powers $P_{\mathrm{d}}^{\mathrm{B}}$ and $P_{\mathrm{u}}^{\mathrm{B}}$ then follow directly from $\mathbf{p}^*$ as shown below, viz. (29), (30).

Proof: In order to arrive at the above closed-form solution, let us first rewrite Constraints C1 and C2 from (22) in terms of the individual UL and DL transmit powers as follows:

$$p_i^{\mathrm{d}} \geq \frac{(2^{\rho_{\mathrm{d}}}-1)\left(\sigma_{\mathrm{n}}^2 + L_i^{\mathrm{Bd}}P_{\mathrm{d}}^{\mathrm{B}} + \sum_{j=1}^{U}L_{ij}^{\mathrm{ud}}p_j^{\mathrm{u}}\right)}{\Lambda_{\mathrm{t}}^{\mathrm{FD}}L_i^{\mathrm{d}}}, \tag{25}$$

$$p_j^{\mathrm{u}} \geq \frac{(2^{\rho_{\mathrm{u}}}-1)\left(\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{AN}}\left[P_{\mathrm{u}}^{\mathrm{B}} + \sum_{i=1}^{D}p_i^{\mathrm{d}}\right]\right)}{\Lambda_{\mathrm{r}}^{\mathrm{FD}}L_j^{\mathrm{u}}}. \tag{26}$$

Minimizing p_i^{d} and p_j^{u} by setting them equal to their lower bounds, we can write $R_i^{\mathrm{d}} = \rho_{\mathrm{d}}\ \forall i$ and $R_j^{\mathrm{u}} = \rho_{\mathrm{u}}\ \forall j$. Hence, the backhauling constraints become $R^{\mathrm{d,B}} \geq D\rho_{\mathrm{d}}$ and $R^{\mathrm{u,B}} \geq U\rho_{\mathrm{u}}$. Utilizing (12) and (13), the following lower bounds for the backhaul-related transmit powers are obtained:

$$P_{\mathrm{d}}^{\mathrm{B}} \geq \gamma_{\mathrm{r}}^{\mathrm{B}}\left[\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{AN}}\left(P_{\mathrm{u}}^{\mathrm{B}} + \sum_{i=1}^{D}p_i^{\mathrm{d}}\right)\right], \tag{27}$$

$$P_{\mathrm{u}}^{\mathrm{B}} \geq \gamma_{\mathrm{t}}^{\mathrm{B}}\left(\sigma_{\mathrm{n}}^2 + \alpha_{\mathrm{BN}}P_{\mathrm{d}}^{\mathrm{B}} + \sum_{j=1}^{U}L_j^{\mathrm{Bu}}p_j^{\mathrm{u}}\right). \tag{28}$$

These transmit powers are also minimized by setting them equal to their lower bounds. Solving then for $P_{\mathrm{d}}^{\mathrm{B}}$ and $P_{\mathrm{u}}^{\mathrm{B}}$ in terms of p_i^{d} and p_j^{u} from (27) and (28), we get:

$$\begin{aligned} P_{\mathrm{d}}^{\mathrm{B}} &= \frac{\alpha_{\mathrm{AN}}\gamma_{\mathrm{r}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}}\sum_{i=1}^{D}p_i^{\mathrm{d}} + \frac{\alpha_{\mathrm{AN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}}\sum_{j=1}^{U}L_j^{\mathrm{Bu}}p_j^{\mathrm{u}} \\ &\quad + \frac{\left(1+\alpha_{\mathrm{AN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\right)\gamma_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{1-\gamma^{\mathrm{B}}}, \end{aligned} \tag{29}$$

$$\begin{aligned} P_{\mathrm{u}}^{\mathrm{B}} &= \frac{\gamma^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}}\sum_{i=1}^{D}p_i^{\mathrm{d}} + \frac{\gamma_{\mathrm{t}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}}\sum_{j=1}^{U}L_j^{\mathrm{Bu}}p_j^{\mathrm{u}} \\ &\quad + \frac{\left(1+\alpha_{\mathrm{BN}}\gamma_{\mathrm{r}}^{\mathrm{B}}\right)\gamma_{\mathrm{t}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{1-\gamma^{\mathrm{B}}}. \end{aligned} \tag{30}$$

Then, by substituting (29) into (25) and (30) into (26), we get the following system of $D+U$ equations with $D+U$ unknown transmit powers:

$$\mathbf{W}_{\mathrm{FD}}\mathbf{p} = \mathbf{v}_{\mathrm{FD}}, \tag{31}$$

where, by denoting a $N \times N$ identity matrix by $\mathbf{I}_N$, $\mathbf{W}_{\mathrm{FD}}$ can be written in blockwise form as

$$\mathbf{W}_{\mathrm{FD}} = \begin{bmatrix} \mathbf{I}_D - \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}\mathbf{q}_{\mathrm{B/d}}\mathbf{1}_D^T}{1-\gamma^{\mathrm{B}}} & -\frac{\gamma_{\mathrm{r}}^{\mathrm{B}}\gamma_{\mathrm{t}}^{\mathrm{B}}\gamma_{\mathrm{d}}\mathbf{q}_{\mathrm{B/d}}\mathbf{L}_{\mathrm{Bu}}^T}{1-\gamma^{\mathrm{B}}} - \frac{\gamma_{\mathrm{d}}\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}}{\alpha_{\mathrm{AN}}} \\ -\frac{\gamma_{\mathrm{u}}\mathbf{q}_{\mathrm{u}}\mathbf{1}_D^T}{1-\gamma^{\mathrm{B}}} & \mathbf{I}_U - \frac{\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}\mathbf{q}_{\mathrm{u}}\mathbf{L}_{\mathrm{Bu}}^T}{1-\gamma^{\mathrm{B}}} \end{bmatrix}, \tag{32}$$

while the vector $\mathbf{v}_{\mathrm{FD}}$ is defined as follows:

$$\mathbf{v}_{\mathrm{FD}} = \begin{bmatrix} \frac{\gamma_{\mathrm{d}}\sigma_{\mathrm{n}}^2}{\alpha_{\mathrm{AN}}}\mathbf{q}_{\mathrm{d}} + \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{1-\gamma^{\mathrm{B}}}\left(\frac{1}{\alpha_{\mathrm{AN}}} + \gamma_{\mathrm{t}}^{\mathrm{B}}\right)\mathbf{q}_{\mathrm{B/d}} \\ \frac{\gamma_{\mathrm{u}}\sigma_{\mathrm{n}}^2}{1-\gamma^{\mathrm{B}}}\left(\frac{1}{\alpha_{\mathrm{AN}}} + \gamma_{\mathrm{t}}^{\mathrm{B}}\right)\mathbf{q}_{\mathrm{u}} \end{bmatrix}. \tag{33}$$

The solution to (31) can then simply be obtained by

$$\mathbf{p}^* = \mathbf{W}_{\mathrm{FD}}^{-1}\mathbf{v}_{\mathrm{FD}}. \tag{34}$$

To express the inverse of the matrix $\mathbf{W}_{\mathrm{FD}}$, let us first write it as a sum of three matrices as follows:

$$\mathbf{W}_{\mathrm{FD}} = \mathbf{I}_{D+U} + \mathbf{F}\mathbf{G} + \mathbf{H}, \tag{35}$$

where

$$\mathbf{F} = \begin{bmatrix} -\frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}}\mathbf{q}_{\mathrm{B/d}} & -\frac{\gamma_{\mathrm{r}}^{\mathrm{B}}\gamma_{\mathrm{t}}^{\mathrm{B}}\gamma_{\mathrm{d}}}{1-\gamma^{\mathrm{B}}}\mathbf{q}_{\mathrm{B/d}} \\ -\frac{\gamma_{\mathrm{u}}}{1-\gamma^{\mathrm{B}}}\mathbf{q}_{\mathrm{u}} & -\frac{\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}}{1-\gamma^{\mathrm{B}}}\mathbf{q}_{\mathrm{u}} \end{bmatrix}$$
$$\mathbf{G} = \begin{bmatrix} \mathbf{1}_D^T & \mathbf{0}_U^T \\ \mathbf{0}_D^T & \mathbf{L}_{\mathrm{Bu}}^T \end{bmatrix}$$
$$\mathbf{H} = \begin{bmatrix} \mathbf{0}_{D\times D} & -\frac{\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}} \\ \mathbf{0}_{U\times D} & \mathbf{0}_{U\times U} \end{bmatrix}.$$

Here, $\mathbf{0}_N$ refers to a column vector consisting of N zeros, while $\mathbf{0}_{M\times N}$ refers to an $M \times N$ matrix consisting of all zeros. Now, the inverse can be written as follows:

$$\begin{aligned}\mathbf{W}_{\mathrm{FD}}^{-1} &= (\mathbf{I}_{D+U} + \mathbf{H} + \mathbf{FG})^{-1} \\ &= (\mathbf{I}_{D+U} + \mathbf{H})^{-1} - (\mathbf{I}_{D+U} + \mathbf{H})^{-1}\mathbf{F} \\ &\times \left(\mathbf{I}_2 + \mathbf{G}\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1}\mathbf{F}\right)^{-1}\mathbf{G}\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1},\end{aligned} \tag{36}$$

where the latter form is obtained using the Kailath Variant [43]. The inverse of the matrix $\mathbf{I}_{D+U} + \mathbf{H}$ can easily be obtained as:

$$\begin{aligned}\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1} &= \begin{bmatrix} \mathbf{I}_{D\times D} & -\frac{\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}} \\ \mathbf{0}_{U\times D} & \mathbf{I}_{U\times U} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \mathbf{I}_{D\times D} & \frac{\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}} \\ \mathbf{0}_{U\times D} & \mathbf{I}_{U\times U} \end{bmatrix}.\end{aligned}$$

Furthermore, since $\mathbf{I}_2 + \mathbf{G}\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1}\mathbf{F}$ is in fact a 2×2 matrix, its inverse can also be calculated in a straightforward manner. In particular, we get

$$\begin{aligned}&\left(\mathbf{I}_2 + \mathbf{G}\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1}\mathbf{F}\right)^{-1} \\ &= \frac{1}{\theta}\begin{bmatrix} 1 - \frac{\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}\mathbf{L}_{\mathrm{Bu}}^T\mathbf{q}_{\mathrm{u}}}{1-\gamma^{\mathrm{B}}} & \frac{\gamma_{\mathrm{r}}^{\mathrm{B}}\gamma_{\mathrm{t}}^{\mathrm{B}}\gamma_{\mathrm{d}}\mathbf{1}_D^T\mathbf{q}_{\mathrm{B/d}}}{1-\gamma^{\mathrm{B}}} + \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}\mathbf{1}_D^T\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}\mathbf{q}_{\mathrm{u}}}{\alpha_{\mathrm{AN}}(1-\gamma^{\mathrm{B}})} \\ \frac{\gamma_{\mathrm{u}}\mathbf{L}_{\mathrm{Bu}}^T\mathbf{q}_{\mathrm{u}}}{1-\gamma^{\mathrm{B}}} & 1 - \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}\mathbf{1}_D^T\mathbf{q}_{\mathrm{B/d}}}{1-\gamma^{\mathrm{B}}} - \frac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}}\mathbf{1}_D^T\mathbf{L}_{\mathrm{ud}}^{\mathrm{d}}\mathbf{q}_{\mathrm{u}}}{\alpha_{\mathrm{AN}}(1-\gamma^{\mathrm{B}})} \end{bmatrix},\end{aligned}$$

where θ is the determinant of the inverted 2×2 matrix, and it is defined as shown in (24).

Having now calculated all the inverses in (36), the optimal transmit powers can be expressed by using only vector/matrix-multiplications as follows:

$$\begin{aligned}\mathbf{p}^* = \Big[&\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1} - \left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1} \\ &\times \mathbf{F}\left(\mathbf{I}_2 + \mathbf{G}\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1}\mathbf{F}\right)^{-1}\mathbf{G}\left(\mathbf{I}_{D+U} + \mathbf{H}\right)^{-1}\Big]\mathbf{v}_{\mathrm{FD}},\end{aligned}$$

which, after substituting the matrices with the corresponding expressions and manipulating the equation, results in (23). These DL and UL transmit powers can then be substituted into (29) and (30) in order to obtain the corresponding optimal backhaul-related transmit powers. ∎

B. Half-Duplex Scheme

Theorem 2: For the half-duplex scheme, the optimal transmit powers in closed form are

$$\mathbf{p}^* = \begin{bmatrix} \mathbf{p}_{\mathrm{d}}^* \\ \mathbf{p}_{\mathrm{u}}^* \end{bmatrix} = \begin{bmatrix} \frac{\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}} - 1\right)\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{t}}^{\mathrm{HD}}}\mathbf{q}_{\mathrm{d}} \\ \frac{\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}} - 1\right)\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{r}}^{\mathrm{HD}}}\mathbf{q}_{\mathrm{u}} \end{bmatrix}, \tag{37}$$

$$P_{\mathrm{d}}^{\mathrm{B}} = \frac{\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}(1-\eta)}} - 1\right)M_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{r}}^{\mathrm{HD}}L_{\mathrm{B}}}, \tag{38}$$

$$P_{\mathrm{u}}^{\mathrm{B}} = \frac{\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}\eta}} - 1\right)M_{\mathrm{t}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{t}}^{\mathrm{HD}}L_{\mathrm{B}}}. \tag{39}$$

The QoS requirements can always be fulfilled and the system is always feasible.

Proof: This analytical solution is again obtained by first rewriting the Constraints C1 and C2 in terms of the DL and UL transmit powers with the help of (14) and (15) as follows:

$$p_i^{\mathrm{d}} \geq \frac{\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}} - 1\right)\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{t}}^{\mathrm{HD}}L_i^{\mathrm{d}}}, \tag{40}$$

$$p_j^{\mathrm{u}} \geq \frac{\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}} - 1\right)\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{r}}^{\mathrm{HD}}L_j^{\mathrm{u}}}. \tag{41}$$

Again, these transmit powers are minimized by setting them equal to the lower bounds, and consequently the backhauling requirements become $R^{\mathrm{d,B}} \geq D\rho_{\mathrm{d}}$ and $R^{\mathrm{u,B}} \geq U\rho_{\mathrm{u}}$. These, together with (16) and (17), yield the following bounds for the backhaul-related transmit powers:

$$P_{\mathrm{d}}^{\mathrm{B}} \geq \frac{\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}(1-\eta)}} - 1\right)M_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{r}}^{\mathrm{HD}}L_{\mathrm{B}}}, \tag{42}$$

$$P_{\mathrm{u}}^{\mathrm{B}} \geq \frac{\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}\eta}} - 1\right)M_{\mathrm{t}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{t}}^{\mathrm{HD}}L_{\mathrm{B}}}, \tag{43}$$

which are also minimized by setting them equal to their respective lower bounds. ∎

Optimizing the Duplexing Parameter for the Half-duplex Scheme: In addition, for the half-duplex scheme, also the duplexing parameter η can be optimized, since it directly affects the overall transmit power. Having solved the optimal transmit powers as shown above, they can be used to formulate the optimization problem in terms of η as follows:

$$\underset{\eta}{\text{minimize}} \quad S^{\mathrm{HD}}(\eta), \quad \text{where } S^{\mathrm{HD}}(\eta) = \mathbf{1}^T\mathbf{p} + P_{\mathrm{d}}^{\mathrm{B}} + P_{\mathrm{u}}^{\mathrm{B}}. \tag{44}$$

Using (37)–(39), the objective function can be written as follows:

$$\begin{aligned}S^{\mathrm{HD}}(\eta) = &\frac{\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}} - 1\right)\sigma_{\mathrm{n}}^2 S_{\mathrm{d}}}{\Lambda_{\mathrm{t}}^{\mathrm{HD}}} + \frac{\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}} - 1\right)\sigma_{\mathrm{n}}^2 S_{\mathrm{u}}}{\Lambda_{\mathrm{r}}^{\mathrm{HD}}} \\ &+ \frac{\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}(1-\eta)}} - 1\right)M_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{r}}^{\mathrm{HD}}L_{\mathrm{B}}} + \frac{\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}\eta}} - 1\right)M_{\mathrm{t}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2}{\Lambda_{\mathrm{t}}^{\mathrm{HD}}L_{\mathrm{B}}},\end{aligned}$$

where $S_{\rm u} = \mathbf{1}_U^T \mathbf{q}_{\rm u}$. It is easy to show that this is a convex function in terms of η, and hence its global minimum is found at the zero-point of its derivative. Hence, (44) is in fact equivalent to solving the following equation:

$$\begin{aligned}&\frac{\rm d}{{\rm d}\eta} S^{\rm HD}(\eta)\\ &= \ln(2)\,\sigma_{\rm n}^2 \left(-\left(\frac{S_{\rm d}\rho_{\rm d}}{\Lambda_{\rm t}^{\rm HD}}\right)\frac{2^{\frac{\rho_{\rm d}}{\eta}}}{\eta^2} + \left(\frac{S_{\rm u}\rho_{\rm u}}{\Lambda_{\rm r}^{\rm HD}}\right)\frac{2^{\frac{\rho_{\rm u}}{1-\eta}}}{(1-\eta)^2} \right.\\ &\left. + \left(\frac{D\rho_{\rm d}}{\Lambda_{\rm r}^{\rm HD} L_{\rm B}}\right)\frac{2^{\frac{D\rho_{\rm d}}{M_{\rm r}^{\rm B}(1-\eta)}}}{(1-\eta)^2} - \left(\frac{U\rho_{\rm u}}{\Lambda_{\rm t}^{\rm HD} L_{\rm B}}\right)\frac{2^{\frac{U\rho_{\rm u}}{M_{\rm t}^{\rm B}\eta}}}{\eta^2}\right) = 0. \quad (45)\end{aligned}$$

This equation does not have a closed-form solution, but it can be easily solved numerically as done in the performance results of Section V when determining the value of the duplexing parameter η.

C. Hybrid Relay Scheme

Theorem 3: The optimal DL and UL transmit powers for the hybrid relay scheme are

$$\mathbf{p}^* = \begin{bmatrix}\mathbf{p}_{\rm d}^*\\ \mathbf{p}_{\rm u}^*\end{bmatrix} = \begin{bmatrix}\frac{\gamma_{\rm d}\sigma_{\rm n}^2}{\alpha_{\rm AN}}\left(\mathbf{q}_{\rm d} + \frac{\gamma_{\rm r}^{\rm B}(1+S_{\rm d}\gamma_{\rm d})}{1-\gamma_{\rm d}\gamma_{\rm r}^{\rm B} S_{\rm B/d}}\mathbf{q}_{\rm B/d}\right)\\ \frac{\sigma_{\rm n}^2\gamma_{\rm u}}{\alpha_{\rm AN}}\left(\frac{1+\alpha_{\rm AN}\gamma_{\rm t}^{\rm B}}{1-\gamma_{\rm u}\gamma_{\rm t}^{\rm B} S_{\rm B/u}}\mathbf{q}_{\rm u}\right)\end{bmatrix}, \quad (46)$$

when each element of $\mathbf{p}^*$ is positive and finite. Otherwise the QoS requirements cannot be fulfilled and the system is infeasible. Here, $\gamma_{\rm d} = \alpha_{\rm AN}\left(\frac{2^{\rho_{\rm d}/\eta}-1}{\Lambda_{\rm t}^{\rm RL}}\right)$, $\gamma_{\rm u} = \alpha_{\rm AN}\left(\frac{2^{\rho_{\rm u}/(1-\eta)}-1}{\Lambda_{\rm r}^{\rm RL}}\right)$, $\gamma_{\rm r}^{\rm B} = M_{\rm r}^{\rm B}\left(\frac{2^{D\rho_{\rm d}/(M_{\rm r}^{\rm B}\eta)}-1}{\Lambda_{\rm r,B}^{\rm RL} L_{\rm B}}\right)$, and $\gamma_{\rm t}^{\rm B} = M_{\rm t}^{\rm B}\left(\frac{2^{U\rho_{\rm u}/(M_{\rm t}^{\rm B}(1-\eta))}-1}{\Lambda_{\rm t,B}^{\rm RL} L_{\rm B}}\right)$. The backhaul-related transmit powers again directly follow from $\mathbf{p}^*$ as shown below, viz. (49), (50) with equalities.

Proof: Following a similar procedure as in the full-duplex and half-duplex schemes, the first step in obtaining the above closed-form solution is rewriting the QoS constraints in (22) as boundaries for the DL and UL transmit powers using (18) and (19) as follows:

$$p_i^{\rm d} \geq \frac{\left(2^{\frac{\rho_{\rm d}}{\eta}} - 1\right)\left(\sigma_{\rm n}^2 + L_i^{\rm Bd} P_{\rm d}^{\rm B}\right)}{\Lambda_{\rm t}^{\rm RL} L_i^{\rm d}}, \quad (47)$$

$$p_j^{\rm u} \geq \frac{\left(2^{\frac{\rho_{\rm u}}{1-\eta}} - 1\right)\left(\sigma_{\rm n}^2 + \alpha_{\rm AN} P_{\rm u}^{\rm B}\right)}{\Lambda_{\rm r}^{\rm RL} L_j^{\rm u}}. \quad (48)$$

Minimizing again these transmit powers by setting them equal to their lower bounds, the self-backhauling constraints become $R^{\rm d,B} \geq D\rho_{\rm d}$ and $R^{\rm u,B} \geq U\rho_{\rm u}$. Hence, by using (20) and (21), we can write:

$$P_{\rm d}^{\rm B} \geq \gamma_{\rm r}^{\rm B}\left(\sigma_{\rm n}^2 + \alpha_{\rm AN}\sum_{i=1}^{D} p_i^{\rm d}\right), \quad (49)$$

$$P_{\rm u}^{\rm B} \geq \gamma_{\rm t}^{\rm B}\left(\sigma_{\rm n}^2 + \sum_{j=1}^{U} L_j^{\rm Bu} p_j^{\rm u}\right). \quad (50)$$

Setting also these backhaul-related transmit powers equal to their lower bounds and substituting them into (47) and (48), we obtain the following expressions for the individual transmit powers:

$$p_i^{\rm d} = \frac{\gamma_{\rm d}\gamma_{\rm r}^{\rm B} L_i^{\rm Bd}}{L_i^{\rm d}}\sum_{k=1}^{D} p_k^{\rm d} + \frac{\sigma_{\rm n}^2\gamma_{\rm d}}{\alpha_{\rm AN} L_i^{\rm d}}\left(1 + L_i^{\rm Bd}\gamma_{\rm r}^{\rm B}\right), \quad (51)$$

$$p_j^{\rm u} = \frac{\gamma_{\rm u}\gamma_{\rm t}^{\rm B}}{L_j^{\rm u}}\sum_{l=1}^{U} L_l^{\rm Bu} p_l^{\rm u} + \frac{\sigma_{\rm n}^2\gamma_{\rm u}}{\alpha_{\rm AN} L_j^{\rm u}}\left(1 + \alpha_{\rm AN}\gamma_{\rm t}^{\rm B}\right). \quad (52)$$

These can easily be rearranged into a system of equations for the unknown DL and UL transmit powers as follows:

$$\mathbf{W}_{\rm RL}\mathbf{p} = \mathbf{v}_{\rm RL}, \quad (53)$$

where $\mathbf{W}_{\rm RL}$ can be written in blockwise form as

$$\mathbf{W}_{\rm RL} = \begin{bmatrix}\mathbf{I}_D - \gamma_{\rm d}\gamma_{\rm r}^{\rm B}\mathbf{q}_{\rm B/d}\mathbf{1}_D^T & \mathbf{0}_{D\times U}\\ \mathbf{0}_{U\times D} & \mathbf{I}_U - \gamma_{\rm u}\gamma_{\rm t}^{\rm B}\mathbf{q}_{\rm u}\mathbf{L}_{\rm Bu}^T\end{bmatrix}, \quad (54)$$

and the vector $\mathbf{v}_{\rm RL}$ is defined as follows:

$$\mathbf{v}_{\rm RL} = \begin{bmatrix}\frac{\sigma_{\rm n}^2\gamma_{\rm d}}{\alpha_{\rm AN}}\left(\mathbf{q}_{\rm d} + \gamma_{\rm r}^{\rm B}\mathbf{q}_{\rm B/d}\right)\\ \frac{\sigma_{\rm n}^2\gamma_{\rm u}}{\alpha_{\rm AN}}\left(1 + \alpha_{\rm AN}\gamma_{\rm t}^{\rm B}\right)\mathbf{q}_{\rm u}\end{bmatrix}. \quad (55)$$

The optimal transmit powers are then obtained similar to the full-duplex scheme, i.e., from

$$\mathbf{p}^* = \mathbf{W}_{\rm RL}^{-1}\mathbf{v}_{\rm RL}, \quad (56)$$

which, due to the block diagonal nature of the matrix $\mathbf{W}_{\rm RL}$, can in fact be solved separately for the DL and UL transmit powers. Hence, the optimal DL transmit powers are as follows:

$$\begin{aligned}\mathbf{p}_{\rm d}^* &= \left(\mathbf{I}_D - \gamma_{\rm d}\gamma_{\rm r}^{\rm B}\mathbf{q}_{\rm B/d}\mathbf{1}_D^T\right)^{-1}\frac{\sigma_{\rm n}^2\gamma_{\rm d}}{\alpha_{\rm AN}}\left(\mathbf{q}_{\rm d} + \gamma_{\rm r}^{\rm B}\mathbf{q}_{\rm B/d}\right)\\ &= \left(\mathbf{I}_D + \frac{\gamma_{\rm d}\gamma_{\rm r}^{\rm B}\mathbf{q}_{\rm B/d}\mathbf{1}_D^T}{1-\gamma_{\rm d}\gamma_{\rm r}^{\rm B}\mathbf{1}_D^T\mathbf{q}_{\rm B/d}}\right)\frac{\sigma_{\rm n}^2\gamma_{\rm d}}{\alpha_{\rm AN}}\left(\mathbf{q}_{\rm d} + \gamma_{\rm r}^{\rm B}\mathbf{q}_{\rm B/d}\right)\\ &= \frac{\sigma_{\rm n}^2\gamma_{\rm d}}{\alpha_{\rm AN}}\left(\mathbf{q}_{\rm d} + \frac{\gamma_{\rm r}^{\rm B}(1+S_{\rm d}\gamma_{\rm d})}{1-\gamma_{\rm d}\gamma_{\rm r}^{\rm B} S_{\rm B/d}}\mathbf{q}_{\rm B/d}\right), \quad (57)\end{aligned}$$

where the matrix inverse has been calculated by using the Sherman-Morrison formula [43]. The optimal UL transmit powers are obtained in an identical manner, and they read:

$$\begin{aligned}\mathbf{p}_{\rm u}^* &= \left(\mathbf{I}_U - \gamma_{\rm u}\gamma_{\rm t}^{\rm B}\mathbf{q}_{\rm u}\mathbf{L}_{\rm Bu}^T\right)^{-1}\frac{\sigma_{\rm n}^2\gamma_{\rm u}}{\alpha_{\rm AN}}\left(1 + \alpha_{\rm AN}\gamma_{\rm t}^{\rm B}\right)\mathbf{q}_{\rm u}\\ &= \frac{\sigma_{\rm n}^2\gamma_{\rm u}}{\alpha_{\rm AN}}\left(\frac{1+\alpha_{\rm AN}\gamma_{\rm t}^{\rm B}}{1-\gamma_{\rm u}\gamma_{\rm t}^{\rm B} S_{\rm B/u}}\mathbf{q}_{\rm u}\right). \quad (58)\end{aligned}$$

The optimal backhaul-related transmit powers can be solved by substituting the optimal DL and UL transmit powers into the expressions in (49) and (50) with equalities. ∎

Optimizing the Duplexing Parameter for the Hybrid Relay Scheme: Similar to the half-duplex scheme, the solution in (56) is for a given duplexing parameter η, and thus the transmit powers of the hybrid relay scheme can be further minimized also with respect to η. This results in the following optimization problem:

$$\underset{\eta}{\text{minimize}} \quad S^{\rm RL}(\eta), \quad \text{where } S^{\rm RL}(\eta) = \mathbf{1}^T\mathbf{p} + P_{\rm d}^{\rm B} + P_{\rm u}^{\rm B}. \quad (59)$$

In order to obtain the expression of the objective function, the overall DL transmit power can first be written as follows, based on (57):

$$\begin{aligned}\mathbf{1}_D^T\mathbf{p}_{\mathrm{d}} &= \frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}\left(\mathbf{1}_D^T\mathbf{q}_{\mathrm{d}} + \frac{\gamma_{\mathrm{r}}^{\mathrm{B}}\left(1+S_{\mathrm{d}}\gamma_{\mathrm{d}}\right)}{1-\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}\mathbf{1}_D^T\mathbf{q}_{\mathrm{B/d}}\right)\\ &= \frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{d}}}{\alpha_{\mathrm{AN}}}\left(\frac{S_{\mathrm{d}}+\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{1-\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}\right).\end{aligned} \tag{60}$$

Following a similar procedure, the sum UL transmit power is obtained using (58) as follows:

$$\mathbf{1}_U^T\mathbf{p}_{\mathrm{u}} = \frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{u}}}{\alpha_{\mathrm{AN}}}\left(\frac{S_{\mathrm{u}}\left(1+\alpha_{\mathrm{AN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\right)}{1-\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\right). \tag{61}$$

Furthermore, in order to express the transmit power used for backhauling the uplink data ($P_{\mathrm{u}}^{\mathrm{B}}$), the term $\mathbf{L}_{\mathrm{Bu}}^T\mathbf{p}_{\mathrm{u}}$ must be calculated. Using an identical procedure as in (61), it can be derived as follows:

$$\mathbf{L}_{\mathrm{Bu}}^T\mathbf{p}_{\mathrm{u}} = \frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{u}}}{\alpha_{\mathrm{AN}}}\left(\frac{S_{\mathrm{B/u}}\left(1+\alpha_{\mathrm{AN}}\gamma_{\mathrm{t}}^{\mathrm{B}}\right)}{1-\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\right). \tag{62}$$

Having obtained the expressions for the sum DL and UL transmit powers, as well as for $\mathbf{L}_{\mathrm{Bu}}^T\mathbf{p}_{\mathrm{u}}$, they can be substituted into the expressions in (49) and (50) to solve the corresponding backhaul-related transmit powers. After this, the objective function can be written as follows:

$$\begin{aligned}S^{\mathrm{RL}}(\eta) = &\frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{d}}\left(\alpha_{\mathrm{AN}}^{-1}+\gamma_{\mathrm{r}}^{\mathrm{B}}\right)\left(S_{\mathrm{d}}+\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}\right)}{1-\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}} + \gamma_{\mathrm{r}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2\\ &+ \frac{\sigma_{\mathrm{n}}^2\gamma_{\mathrm{u}}\left(\alpha_{\mathrm{AN}}^{-1}+\gamma_{\mathrm{t}}^{\mathrm{B}}\right)\left(S_{\mathrm{u}}+\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}\right)}{1-\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}} + \gamma_{\mathrm{t}}^{\mathrm{B}}\sigma_{\mathrm{n}}^2,\end{aligned} \tag{63}$$

where the duplexing parameter η is contained in the terms γ_{d}, γ_{u}, $\gamma_{\mathrm{r}}^{\mathrm{B}}$, and $\gamma_{\mathrm{t}}^{\mathrm{B}}$, as defined earlier. As is shown in Section IV below where the feasibility of the hybrid relay scheme is discussed in more detail, the system is in fact feasible when $1-\gamma_{\mathrm{d}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}} > 0$ and $1-\gamma_{\mathrm{u}}^{\mathrm{B}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}} > 0$. Solving these inequalities in terms of η results in an open interval within which the minimum point is located, and it can be easily observed that the function $S^{\mathrm{RL}}(\eta)$ is also continuous within this interval.

As there is no closed-form solution for the optimal η, in the forthcoming numerical results the optimal duplexing parameter is determined by numerically optimizing (63) over the open interval defined by the feasibility conditions. This can be done by utilizing any one-dimensional optimization procedure, and the hereby obtained optimal value of η is then used when evaluating the minimum transmit powers of the hybrid relay scheme.

IV. Feasibility Analysis of Full-Duplex and Hybrid Relay Schemes

The feasibility of the considered communication schemes can be determined by investigating the resulting required transmit powers. In particular, if their values are positive and finite, the system is capable of fulfilling the QoS requirements, while infinite or negative transmit powers in the above theorems naturally indicate that the required data rates cannot be achieved. This stems from the physical interpretation of a transmit power, which obviously cannot be negative.

The half-duplex scheme does not suffer from any interference sources, and hence it is feasible under all circumstances. In other words, it can fulfill any QoS requirements with appropriately high transmit powers. However, both the full-duplex and hybrid relay schemes have various interference sources, which result in a fundamental upper bound for the achievable data rates. We refer to this as the feasibility boundary, since it determines whether the whole system is feasible in the first place. Essentially, this means that the full-duplex and hybrid relay schemes have an upper bound for the DL and/or UL data rates, which can be expressed as follows:

$$R^{\mathrm{x}} \leq R_{\max}^{\mathrm{x}}, \quad \forall p_1^{\mathrm{d}},\ldots,p_D^{\mathrm{d}},p_1^{\mathrm{u}},\ldots,p_U^{\mathrm{u}},P_{\mathrm{d}}^{\mathrm{B}},P_{\mathrm{u}}^{\mathrm{B}} \geq 0,$$

where $x = d$ and/or $x = u$. This means that, if the DL/UL data rate requirement is higher than $R_{\max}^{\mathrm{x}}$, the QoS requirements cannot be fulfilled, and consequently the system is infeasible. Note that essentially this type of a feasibility analysis considers a case where all the transmit powers tend towards infinity, meaning that the derived boundary conditions are very fundamental in nature. Hence, the corresponding feasibility limits for restricted transmit powers are somewhat stricter.

A. Feasibility of the Full-duplex Scheme

Theorem 4: The feasibility condition of the full-duplex scheme can be expressed as follows:

$$\begin{cases}\dfrac{\gamma_{\mathrm{d}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{1-\gamma^{\mathrm{B}}} + \dfrac{\gamma_{\mathrm{u}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}{1-\gamma^{\mathrm{B}}} + \dfrac{\gamma_{\mathrm{d}}\gamma_{\mathrm{u}}S_{\mathrm{ud}}}{\alpha_{\mathrm{AN}}\left(1-\gamma^{\mathrm{B}}\right)} < 1,\\ \gamma_{\mathrm{B}} < 1,\end{cases} \tag{64}$$

where the first condition is simply $\theta > 0$ rewritten in a slightly different form.

Proof: These feasibility conditions stem from the fact that all the transmit powers in (23) are positive and finite under these conditions. In particular, if $\gamma_{\mathrm{B}} < 1$, all the terms in (23), apart from θ, are always positive. Then, when also $\theta > 0$, all the transmit powers are clearly positive and the system is feasible. It is also evident from (23) that $\theta < 0$ and $\gamma^{\mathrm{B}} < 1$ result in at least the UL transmit powers being negative, while $\gamma^{\mathrm{B}} > 1$ results in $\theta > 0$, meaning that the UL transmit powers are negative also in this case. This proves that the system is infeasible if and only if the conditions in (64) do not hold. ■

Corollary 1: For any typical system parameters, the term $\alpha_{\mathrm{AN}}\alpha_{\mathrm{BN}}$ is extremely small, meaning that usually $\theta > 0$ is a sufficient feasibility condition since $\gamma^{\mathrm{B}} \ll 1$. In fact, it can be assumed with high accuracy that $\gamma^{\mathrm{B}} = 0$, and the feasibility condition in terms of the physical system parameters can consequently be approximated as:

$$\begin{aligned}\alpha_{\mathrm{AN}}\Bigg(&\frac{(2^{\rho_{\mathrm{u}}}-1)(2^{\rho_{\mathrm{d}}}-1)S_{\mathrm{ud}}}{\Lambda_{\mathrm{t}}^{\mathrm{FD}}\Lambda_{\mathrm{r}}^{\mathrm{FD}}} + \frac{(2^{\rho_{\mathrm{u}}}-1)(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}}-1)M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}{\Lambda_{\mathrm{t}}^{\mathrm{FD}}\Lambda_{\mathrm{r}}^{\mathrm{FD}}L_{\mathrm{B}}}\\ &+ \frac{(2^{\rho_{\mathrm{d}}}-1)(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}}-1)M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{\Lambda_{\mathrm{t}}^{\mathrm{FD}}\Lambda_{\mathrm{r}}^{\mathrm{FD}}L_{\mathrm{B}}}\Bigg) < 1.\end{aligned} \tag{65}$$

Since (65) is clearly a monotonic function of ρ_{d}, the maximum supported DL data rate requirement is obtained

by solving the root of the above expression with respect to ρ_{d}. Due to the multiplications of the exponential rate terms, there is no closed-form solution for the root, even if assuming that $2^{\rho_{\mathrm{d}}}-1\approx 2^{\rho_{\mathrm{d}}}$ and $2^{D\rho_{\mathrm{d}}/M_{\mathrm{r}}^{\mathrm{B}}}-1\approx 2^{D\rho_{\mathrm{d}}/M_{\mathrm{r}}^{\mathrm{B}}}$. Hence, the highest feasible DL data rate requirement of the full-duplex scheme is obtained by solving the root numerically.

On the other hand, when considering the required amount of SI cancellation to make the system feasible, a closed-form solution can be easily obtained from (65). In particular, the minimum amount of required SI cancellation in decibels is as follows:

$$\begin{aligned}\alpha_{\mathrm{AN}}^{\mathrm{dB}} < {} & 10\log_{10}\left(\Lambda_{\mathrm{t}}^{\mathrm{FD}}\Lambda_{\mathrm{r}}^{\mathrm{FD}}L_{\mathrm{B}}\right)\\ & -10\log_{10}\Bigg[\left(2^{\rho_{\mathrm{u}}}-1\right)\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}}-1\right)M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}\\ & \quad +\left(2^{\rho_{\mathrm{d}}}-1\right)\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}}-1\right)M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}\\ & \quad +\left(2^{\rho_{\mathrm{u}}}-1\right)\left(2^{\rho_{\mathrm{d}}}-1\right)L_{\mathrm{B}}S_{\mathrm{ud}}\Bigg]. \end{aligned} \tag{66}$$

In the numerical results, these feasibility boundaries obtained from the simplified expression in Corollary 1 are compared to the exact solutions defined in Theorem 4. The approximated boundaries are shown to be highly accurate, which means that Corollary 1 can be used to obtain reliable information regarding the feasibility of a system utilizing the full-duplex scheme.

B. Feasibility of the Hybrid Relay Scheme

Theorem 5: The hybrid relay scheme is feasible under the following conditions:

$$\begin{cases} \dfrac{\alpha_{\mathrm{AN}}\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}}-1\right)\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}\eta}}-1\right)M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{\Lambda_{\mathrm{r,B}}^{\mathrm{RL}}\Lambda_{\mathrm{t}}^{\mathrm{RL}}L_{\mathrm{B}}}<1, \\ \dfrac{\alpha_{\mathrm{AN}}\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}}-1\right)\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}(1-\eta)}}-1\right)M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}{\Lambda_{\mathrm{t,B}}^{\mathrm{RL}}\Lambda_{\mathrm{r}}^{\mathrm{RL}}L_{\mathrm{B}}}<1, \\ 0<\eta<1. \end{cases} \tag{67}$$

Proof: These conditions are obtained by observing from (46) that all the transmit powers are positive and finite when $1-\gamma_{\mathrm{d}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}>0$ and $1-\gamma_{\mathrm{u}}^{\mathrm{B}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}>0$, because all the variables themselves are positive. Furthermore, since the sum DL and UL transmit powers in (60) and (61) are negative when $1-\gamma_{\mathrm{d}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}<0$ and $1-\gamma_{\mathrm{u}}^{\mathrm{B}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}<0$, (67) represents indeed the exact feasibility condition. ∎

When optimizing the duplexing parameter η, these conditions can be used to determine its upper and lower bound. In particular, it can easily be shown that the first condition is monotonically decreasing with respect to η, while the second condition is monotonically increasing. Hence, the first inequality results in a lower bound for η, while the second inequality defines its upper bound. The system is then feasible if there exists a value for η which fulfills all of these inequalities.

Since it is not possible to obtain closed-form solutions for the upper and lower boundaries of η using the exact form of (67), the problem can be made analytically tractable by assuming that $\left(2^{\frac{\rho_{\mathrm{d}}}{\eta}}-1\right)\approx 2^{\frac{\rho_{\mathrm{d}}}{\eta}}$, $\left(2^{\frac{\rho_{\mathrm{u}}}{1-\eta}}-1\right)\approx 2^{\frac{\rho_{\mathrm{u}}}{1-\eta}}$, $\left(2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}\eta}}-1\right)\approx 2^{\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}\eta}}$, and $\left(2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}(1-\eta)}}-1\right)\approx 2^{\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}(1-\eta)}}$. This approximation is rather accurate with any reasonable rate requirements, and it represents a pessimistic estimate of the feasibility boundary, which is asymptotically approaching the true boundary when $\rho_{\mathrm{d}},\rho_{\mathrm{u}}\to\infty$. Now, the boundaries for η can be expressed as follows:

$$\frac{\rho_{\mathrm{d}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}}{\log_2\left(\frac{\Lambda_{\mathrm{r,B}}^{\mathrm{RL}}\Lambda_{\mathrm{t}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}\right)}<\eta<1-\frac{\rho_{\mathrm{u}}+\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}}{\log_2\left(\frac{\Lambda_{\mathrm{r}}^{\mathrm{RL}}\Lambda_{\mathrm{t,B}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\right)}.$$

Note that we have assumed here that $\frac{\Lambda_{\mathrm{r,B}}^{\mathrm{RL}}\Lambda_{\mathrm{t}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}>1$ and $\frac{\Lambda_{\mathrm{r}}^{\mathrm{RL}}\Lambda_{\mathrm{t,B}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}>1$, since this ensures that the third condition, i.e., $0<\eta<1$, is fulfilled. Because these inequalities can be expected to hold when considering any realistic system parameters, they are not explicitly analyzed in this article.

Corollary 2: Noting that, for a feasible system, the lower bound of η must be strictly less than its upper bound, an approximative feasibility condition for the hybrid relay scheme can be expressed as

$$\frac{\rho_{\mathrm{d}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}}{\log_2\left(\frac{\Lambda_{\mathrm{r,B}}^{\mathrm{RL}}\Lambda_{\mathrm{t}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}\right)}+\frac{\rho_{\mathrm{u}}+\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}}{\log_2\left(\frac{\Lambda_{\mathrm{r}}^{\mathrm{RL}}\Lambda_{\mathrm{t,B}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\right)}<1. \tag{68}$$

If the lower bound is equal to the upper bound, this means that $1-\gamma_{\mathrm{d}}^{\mathrm{B}}\gamma_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}\approx 0$ and $1-\gamma_{\mathrm{u}}^{\mathrm{B}}\gamma_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}\approx 0$, indicating that the required transmit powers tend to infinity, and hence this condition represents the feasibility boundary.

Using (68), we can easily derive the boundary for the DL data rate requirement with respect to the other system parameters, and it is as follows:

$$\rho_{\mathrm{d}}<\frac{\log_2\left(\frac{\Lambda_{\mathrm{r,B}}^{\mathrm{RL}}\Lambda_{\mathrm{t}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}\right)}{1+\frac{D}{M_{\mathrm{r}}^{\mathrm{B}}}}\left(1-\frac{\rho_{\mathrm{u}}+\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}}{\log_2\left(\frac{\Lambda_{\mathrm{r}}^{\mathrm{RL}}\Lambda_{\mathrm{t,B}}^{\mathrm{RL}}L_{\mathrm{B}}}{\alpha_{\mathrm{AN}}M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\right)}\right). \tag{69}$$

The minimum requirement for SI cancellation in the AN can also be written in closed form using (68). Expressing α_{AN} in decibels, it reads as follows:

$$\begin{aligned}\alpha_{\mathrm{AN}}^{\mathrm{dB}} < {} & 5\log_{10}\left(\frac{\Lambda_{\mathrm{r}}^{\mathrm{RL}}\Lambda_{\mathrm{t}}^{\mathrm{RL}}\Lambda_{\mathrm{t,B}}^{\mathrm{RL}}\Lambda_{\mathrm{r,B}}^{\mathrm{RL}}L_{\mathrm{B}}^2}{M_{\mathrm{r}}^{\mathrm{B}}M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/d}}S_{\mathrm{B/u}}}\right)\\ & -\frac{5}{\log_2(10)}\Bigg[\rho_{\mathrm{d}}+\rho_{\mathrm{u}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}+\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}\\ & +\Bigg[\left(\rho_{\mathrm{d}}-\rho_{\mathrm{u}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}-\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}\right.\\ & \left.+\log_2\left(\frac{\Lambda_{\mathrm{t,B}}^{\mathrm{RL}}\Lambda_{\mathrm{r}}^{\mathrm{RL}}M_{\mathrm{r}}^{\mathrm{B}}S_{\mathrm{B/d}}}{\Lambda_{\mathrm{t}}^{\mathrm{RL}}\Lambda_{\mathrm{r,B}}^{\mathrm{RL}}M_{\mathrm{t}}^{\mathrm{B}}S_{\mathrm{B/u}}}\right)\right)^2\\ & +4\left(\rho_{\mathrm{d}}+\frac{D\rho_{\mathrm{d}}}{M_{\mathrm{r}}^{\mathrm{B}}}\right)\left(\rho_{\mathrm{u}}+\frac{U\rho_{\mathrm{u}}}{M_{\mathrm{t}}^{\mathrm{B}}}\right)\Bigg]^{1/2}\Bigg]. \end{aligned} \tag{70}$$

Note that solving (68) for α_{AN} requires solving the roots of a 2nd-degree polynomial, but one of the two solutions can

TABLE II: The essential default system parameters. Many of the parameter values are also varied in the evaluations.

Parameter	**Value**
No. of AN TX/RX antennas ($N_{\mathrm{t}}/N_{\mathrm{r}}$)	200/100
No. of DL and UL UEs ($D = U$)	10
No. of DL/UL backhaul streams ($M_{\mathrm{r}}^{\mathrm{B}}/M_{\mathrm{t}}^{\mathrm{B}}$)	12/6
Receiver noise floor (σ_{n}^2)	-90 dBm
SI cancellation in the AN/BN ($\alpha_{\mathrm{AN}}/\alpha_{\mathrm{BN}}$)	$-120/-120$ dB
Per-UE DL/UL rate requirement ($\rho_{\mathrm{d}}/\rho_{\mathrm{u}}$)	8/2 bps/Hz
Cell radius	50 m
Distance between the AN and the BN	75 m
No. of Monte Carlo simulation runs	10 000

easily be shown to result in the duplexing parameter being outside the open interval $(0, 1)$. Hence, there is only one valid solution for the inequality. In Section V, the above feasibility boundaries given by Corollary 2 are shown to be very close to the exact feasibility boundaries given by Theorem 5. Hence, these approximative closed-form boundaries provide highly accurate results when determining the feasibility of the hybrid relay scheme.

V. Numerical Results

Next, the proposed system is numerically evaluated with Monte Carlo simulations. In particular, we consider a cell of a given size where the specified amount of DL and UL UEs are randomly positioned. By calculating the optimal transmit powers and the feasibility conditions for a large number of random positions, the cumulative distribution functions (CDFs) of the corresponding quantities can then be obtained. The default system parameters, which are used unless otherwise mentioned, are shown in Table II. The path losses between the different parties are calculated based on the distances in each random realization, using the measurement-based path loss model for a center frequency of 3.5 GHz presented in [44] to reflect a concrete practical example; the line-of-sight (LOS) model is adopted for the link between the AN and the BN, while the non-line-of-sight (NLOS) model is adopted in all the other cases. To ensure a practical system, the scheduled DL and UL UEs are chosen from the opposite sides of the cell, which results in a smaller level of UL-to-DL IUI [41]. The UEs can then alternate between DL and UL modes at regular intervals, by which each UE gets served both in the DL and in the UL, regardless of their position in the cell. Furthermore, in order to facilitate a fair comparison between the different schemes, in the forthcoming figures the transmit powers of the half-duplex and hybrid relay scheme are weighted by the proportion of time spent in the corresponding time slot (determined by the duplexing parameter η). For brevity, the full-duplex, half-duplex and hybrid relay schemes are referred to as FD, HD, and RL, respectively, in all the figures.

A. Feasibility

In order to first analyze the feasibility of the full-duplex and the hybrid relay schemes, Fig. 3 shows the CDFs of the SI cancellation performance required in the AN to make

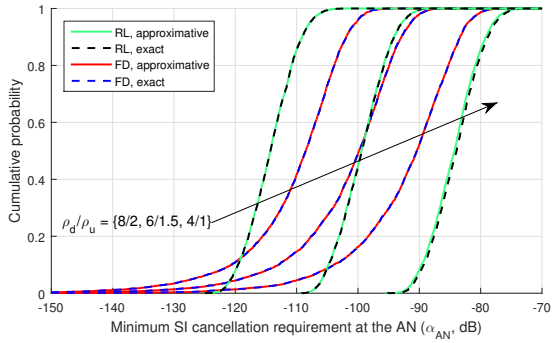


Fig. 3: CDFs of the minimum SI cancellation requirement in the AN in the full-duplex and hybrid relay schemes, shown for different DL/UL data rate requirements.

the system feasible. The figure shows both the approximated closed-form solutions given in (66) and (70) as well as the exact solutions obtained from (64) and (67). Firstly, it can be observed that the approximated feasibility boundaries match the exact boundaries very closely, indicating that the approximations do not compromise the accuracy of the derived equations. Furthermore, Fig. 3 indicates that the required AN SI cancellation performance of the full-duplex scheme is less affected by the data rate requirements than that of the hybrid relay scheme. In particular, with the highest data rate requirements, the full-duplex scheme is feasible with lower SI cancellation performance than the hybrid relay scheme, while the opposite is true for the lowest considered data rate requirements. In the latter case, the hybrid relay scheme benefits from the fact that it only needs to transmit to the UEs *or* to the BN, unlike the full-duplex scheme which must transmit everything at the same time. This results in less stringent SI cancellation requirements. However, with the higher data rate requirements, this benefit is overshadowed by the need to perform time-division duplexing.

Another perspective into the feasibility is the highest supported DL data rate requirement. The corresponding CDFs are plotted in Fig. 4, which again show the approximated boundaries given by (65) and (69), alongside with the exact feasibility boundaries obtained from (64) and (67). Also now, the approximated feasibility boundaries are essentially similar to the exact boundaries, further confirming their accuracy under the studied conditions. It can also be concluded that the full-duplex scheme can support a higher DL data rate requirement with all the considered UL data rate requirements. However, it should be noted that there is more uncertainty regarding the maximum supported DL data rate requirement in the full-duplex scheme, since the slope of the CDF is lower than in the hybrid relay scheme. Hence, even though the full-duplex scheme supports a higher median DL data rate requirement, there is a higher probability that it cannot fulfill that for different randomly chosen UE positions in the network. This indicates that there is a trade-off between the maximum performance and robustness when comparing the full-duplex and hybrid relay schemes.

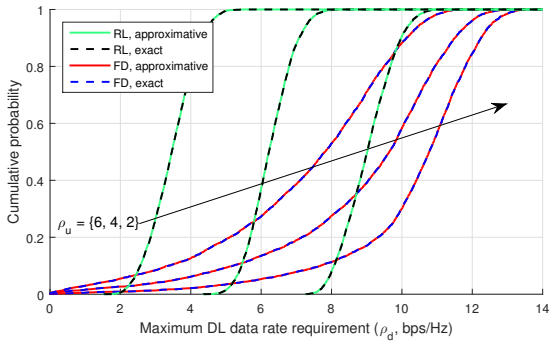


Fig. 4: CDFs of the maximum supported DL data rate requirement in the full-duplex and hybrid relay schemes, shown for different UL data rate requirements.

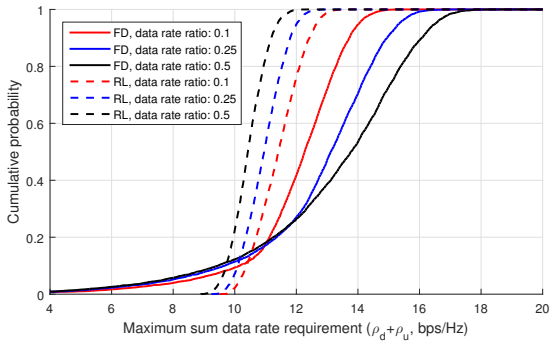


Fig. 5: CDFs of the maximum supported sum data rate requirement in the full-duplex and hybrid relay schemes under different fixed UL/DL data rate ratios.

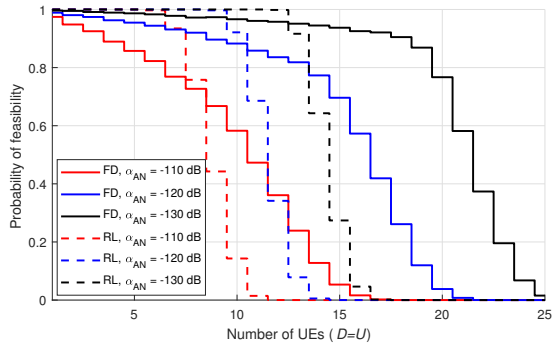


Fig. 6: The probability of feasibility with respect to the number of UEs ($U = D$) in the full-duplex and hybrid relay schemes, shown for different AN SI cancellation performances.

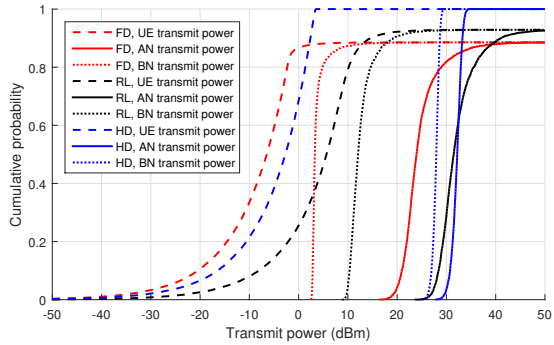


Fig. 7: CDFs of the transmit powers of the individual parties with the default system parameters.

The maximum supported sum data rate requirement is then analyzed in Fig. 5. There, the CDFs of the feasibility boundary are shown under a scenario where the ratio between the UL and DL data rates is fixed, that is, $\rho_{\mathrm{u}}/\rho_{\mathrm{d}} = c$ for some constant c. In this case, the CDFs are only shown for the approximated equations in order to make the figure more readable. In general, the full-duplex scheme supports also a higher median sum data rate requirement, although the uncertainty in the supported data rate requirement is again somewhat higher than in the hybrid relay scheme.

It can also be observed from Fig. 5 that the hybrid relay scheme supports higher sum data rate requirements with lower data rate ratios. This stems from the system parameters having been chosen to support a higher DL data rate ($M_{\mathrm{r}}^{\mathrm{B}} > M_{\mathrm{t}}^{\mathrm{B}}$), which results in the hybrid relay scheme benefiting from a DL-oriented data rate distribution.[1] On the other hand, the full-duplex scheme seems to be better suited for a more even distribution of the DL and UL data rate requirements, which is evident from Fig. 5 when investigating the median values of the highest supported sum-rate requirements. This is due to the more symmetric nature of the full-duplex scheme since it has less options for dividing the resources between UL and DL. Hence, unlike the hybrid relay scheme, which has the benefit of a duplexing parameter, the full-duplex scheme requires a more even data rate distribution to support the highest sum-rates.

Finally, Fig. 6 shows the probability of feasibility with respect to the number of UEs in the full-duplex and hybrid relay schemes, assuming $D = U$. The probabilities have been obtained by evaluating the approximated feasibility boundaries in (65) and (68) for different numbers of UEs. Firstly, it can be observed from Fig. 6 that the full-duplex scheme can in general fulfill the QoS requirements for a larger number of randomly positioned UEs, especially when the AN is capable of efficient SI cancellation. With the lower AN SI cancellation performances, the hybrid relay scheme is more evenly matched with the full-duplex scheme, being again the more robust option in terms of fulfilling the QoS requirements. Namely, while the full-duplex scheme can in general support a larger number of UEs, the slope of the probability curve is steeper with the hybrid relay scheme, indicating that the latter is the more predictable option when there is a moderate number of UEs in the cell. This somewhat resembles the behaviour of the maximum supported DL data rate requirements in Fig. 4. Nevertheless, with sufficiently high AN SI cancellation performance, the full-duplex scheme is clearly the superior option with regard to the number of supported UEs.

[1]Note that the DL data rate requirements can be expected to be higher also in practical networks [25].

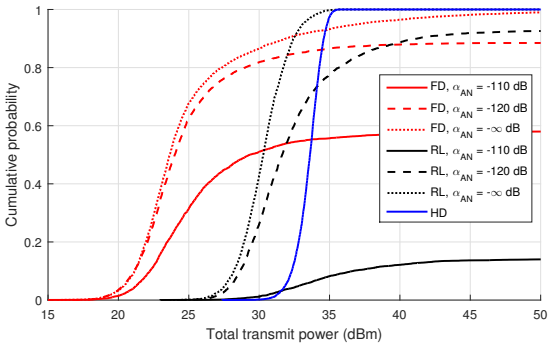


Fig. 8: CDFs of the total used transmit power of each scheme, shown for different values of SI cancellation.

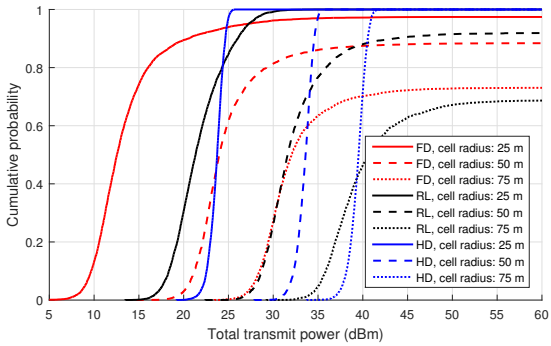


Fig. 9: CDFs of the total used transmit power of each scheme, shown for different cell radii. The distance between the AN and the BN is retained at $\frac{3}{2}$ times the cell radius.

B. Transmit Powers

To then investigate the transmit power efficiency of the different communication schemes, the CDFs of the transmit powers of the AN, each individual UE, and the BN are first shown in Fig. 7 using the default system parameters. It can be observed that the full-duplex scheme obtains the lowest transmit powers for all the communicating parties. However, the downside of the full-duplex scheme is its inability to fulfill the QoS requirements in some cases, evidenced by the fact that the CDFs saturate to a value below 1. These cases represent a situation where the feasibility conditions in (64) are not fulfilled, and therefore the highest value of the CDF is in fact the probability of feasibility of the corresponding system, illustrated also in Fig. 6 with respect to the number of UEs. This deduction is further confirmed by Fig. 3, which shows that the SI cancellation requirement is indeed more than -120 dB in some cases when $\rho_{\mathrm{d}} = 8$ and $\rho_{\mathrm{u}} = 2$.

From the perspective of the overall transmit power consumption, the hybrid relay scheme is then the next best option, while the half-duplex scheme outperforms the hybrid relay scheme in terms of minimizing the UE transmit powers. The reason for this stems from the fact that in the half-duplex scheme the UE transmissions occur in the same time slot where the DL data is backhauled. Due to the higher DL data rate requirements, this results in a somewhat longer time slot for the UE transmissions, allowing for a lower UE transmit power. Note that this occurs due to the optimal duplexing parameter being chosen by minimizing the *total transmit power*. A different outcome would be obtained if a UE-transmit-power-minimizing duplexing parameter was used. What is more, the hybrid relay scheme also suffers from the inability to fulfill the QoS requirements under some circumstances, similar to the full-duplex scheme.

To observe the effect of the SI cancellation capability of the AN, Fig. 8 shows then the CDFs of the total transmit power of the whole radio access system for different values of α_{AN}. Again, the transmit power usage of the full-duplex scheme is significantly lower than that of the other schemes, regardless of the AN SI cancellation performance. However, with the lower values of α_{AN}, the probability of fulfilling the QoS requirements with the full-duplex scheme drops rather low. This is also evident from Fig. 3, where it is clearly seen that the SI cancellation requirement is beyond -110 dB with a large probability when $\rho_{\mathrm{d}} = 8$ and $\rho_{\mathrm{u}} = 2$. Hence, the lower probability of feasibility is the cost of the low transmit power consumption.

The hybrid relay scheme also outperforms the half-duplex scheme when α_{AN} is -120 dB or better, while it performs very poorly with the lowest considered AN SI cancellation performance. This is explained by the CDF of the SI cancellation requirement shown in Fig. 3, which indicates that the SI cancellation requirement of the hybrid relay scheme is in the majority of the cases more than -110 dB. Still, even with $\alpha_{\mathrm{AN}} = -120$ dB, the probability of the hybrid relay scheme having to use more power for the transmissions than the half-duplex scheme is rather high, suggesting that it requires relatively high SI cancellation performance in the AN in order to be a viable option.

To investigate the effect of the cell size on the different schemes, Fig. 9 shows the CDFs of the total transmit power for different cell radii. Again, for all considered cell sizes, the full-duplex scheme is the most power-efficient option, while the hybrid relay scheme and the half-duplex scheme are quite closely matched. Especially with the larger cell sizes, their median transmit power usages are nearly the same. However, the hybrid relay scheme again suffers from the fact that it cannot fulfill the QoS requirements for some UE positions and thus, regardless of the higher median power, the half-duplex scheme might be the more favorable option of these two.

On a more general note, the cell size has a rather significant impact on the required transmit power, as can be expected. For instance, the total median transmit power of the full-duplex scheme is increased by almost 20 dB when the cell radius is increased from 25 m to 75 m. Moreover, with the highest considered cell radius of 75 m, the full-duplex and the hybrid relay schemes cannot fulfill the QoS requirements for a significant portion of the UE positions. Hence, it can be concluded that especially the schemes utilizing inband full-duplex communications are best suited for relatively small cells.

VI. Conclusion

In this paper, we investigated a self-backhauling inband full-duplex access node with large antenna arrays, which can use the same time-frequency resource for serving the mobile users as well as for backhauling, thereby significantly reducing the cost of deployment in ultra-dense networks. Three different communication schemes for the access node were analyzed: a purely full-duplex scheme, a purely half-duplex scheme, and a hybrid scheme where the access node acts as a one-directional full-duplex relay. Especially, we derived the optimal transmit powers for the different communication schemes in closed form when a Quality-of-Service (QoS) requirement for each mobile user is given. In this work, QoS was defined as a minimum achievable data rate. In addition, we showed that the QoS requirements cannot always be achieved when using a full-duplex-capable access node, expressing this feasibility condition also in closed form. Evaluating then the transmit powers and feasibility conditions with realistic system parameter values, it was observed that having a purely full-duplex access node provides usually the lowest transmit powers for all communicating parties. However, the downside of the purely full-duplex scheme is its inability to fulfill the QoS requirements under some circumstances, characterized by the closed-form feasibility conditions. The numerical results also indicated that utilizing a self-backhauling full-duplex access node is best suited for relatively small cells.

References

[1] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.

[2] B. Day, A. Margetts, D. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702–3713, Jul. 2012.

[3] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.

[4] S. Goyal, P. Liu, S. S. Panwar, R. A. Difazio, R. Yang, and E. Bala, "Full duplex cellular systems: will doubling interference prevent doubling capacity?" *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 121–127, May 2015.

[5] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2011, pp. 301–312.

[6] M. Heino et al., "Recent advances in antenna design and interference cancellation algorithms for in-band full-duplex relays," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 91–101, May 2015.

[7] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. SIGCOMM'13*, Aug. 2013, pp. 375–386.

[8] D. Korpi, J. Tamminen, M. Turunen, T. Huusari, Y.-S. Choi, L. Anttila, S. Talwar, and M. Valkama, "Full-duplex mobile device: Pushing the limits," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 80–87, Sep. 2016.

[9] M. Mohammadi, H. A. Suraweera, Y. Cao, I. Krikidis, and C. Tellambura, "Full-duplex radio for uplink/downlink wireless access with spatially random nodes," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 5250–5266, Dec. 2015.

[10] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[11] T. Riihonen, M. Vehkaperä, and R. Wichman, "Large-system analysis of rate regions in bidirectional full-duplex MIMO link: Suppression versus cancellation," in *Proc. 47th Annu. Conf. Inform. Sci. Sys.*, Mar. 2013, pp. 1–6.

[12] D. Korpi, T. Riihonen, and M. Valkama, "Achievable rate regions and self-interference channel estimation in hybrid full-duplex/half-duplex radio links," in *Proc. 47th Annu. Conf. Inform. Sci. Sys.*, Mar. 2015, pp. 1–6.

[13] D. Korpi, T. Riihonen, K. Haneda, K. Yamamoto, and M. Valkama, "Achievable transmission rates and self-interference channel estimation in hybrid full-duplex/half-duplex MIMO relaying," in *Proc. 82nd IEEE Veh. Technol. Conf.*, Sep. 2015.

[14] V. Aggarwal, M. Duarte, A. Sabharwal, and N. Shankaranarayanan, "Full- or half-duplex? A capacity analysis with bounded radio resources," in *Proc. IEEE Inform. Theory Workshop*, Sep. 2012, pp. 207–211.

[15] C. Psomas, M. Mohammadi, I. Krikidis, and H. A. Suraweera, "Impact of directionality on interference mitigation in full-duplex cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 487–502, Jan. 2017.

[16] L. Anttila, D. Korpi, V. Syrjälä, and M. Valkama, "Cancellation of power amplifier induced nonlinear self-interference in full-duplex transceivers," in *Proc. 47th Asilomar Conf. Signals Syst. Comput.*, Nov. 2013, pp. 1193–1198.

[17] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.

[18] E. Ahmed, A. M. Eltawil, and A. Sabharwal, "Self-interference cancellation with nonlinear distortion suppression for full-duplex systems," in *Proc. 47th Asilomar Conf. Signals Syst. Comput.*, Nov. 2013, pp. 1199–1203.

[19] B. Kaufman, J. Lilleberg, and B. Aazhang, "An analog baseband approach for designing full-duplex radios," in *Proc. 47th Asilomar Conf. Signals Syst. Comput.*, Nov. 2013, pp. 987–991.

[20] Y. Liu, X. Quan, W. Pan, S. Shao, and Y. Tang, "Nonlinear distortion suppression for active analog self-interference cancellers in full duplex wireless communication," in *Proc. Globecom Workshops*, Dec. 2014, pp. 948–953.

[21] J. Tamminen, M. Turunen, D. Korpi, T. Huusari, Y.-S. Choi, S. Talwar, and M. Valkama, "Digitally-controlled RF self-interference canceller for full-duplex radios," in *Proc. Europ. Signal Process. Conf.*, Aug. 2016, pp. 783–787.

[22] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, Oct. 2014.

[23] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering full-duplex wireless communication by exploiting directional diversity," in *Proc. 45th Asilomar Conf. Signals Syst. Comput.*, Nov. 2011, pp. 2002–2006.

[24] D. Korpi, T. Riihonen, and M. Valkama, "Inband full-duplex radio access system with self-backhauling: Transmit power minimization under QoS requirements," in *Proc. 42nd Int. Conf. Acoust. Speech Signal Process.*, Mar. 2017.

[25] Nokia Solutions and Networks, "TD-LTE frame configuration primer," Nov. 2013, white paper.

[26] E. Everett and A. Sabharwal, "Spatial degrees-of-freedom in large-array full-duplex: the impact of backscattering," *EURASIP J. Wireless Commun. Netw.*, vol. 2016, no. 1, p. 286, Dec. 2016.

[27] I. Harjula, R. Wichman, K. Pajukoski, E. Lähetkangas, E. Tiirola, and O. Tirkkonen, "Full duplex relaying for local area," in *Proc. 24th Annu. IEEE Int. Symp. Personal, Indoor, and Mobile Radio Commun.*, Sep. 2013, pp. 2684–2688.

[28] R.-A. Pitaval, O. Tirkkonen, R. Wichman, K. Pajukoski, E. Lähetkangas, and E. Tiirola, "Full-duplex self-backhauling for small-cell 5G networks," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 83–89, Oct. 2015.

[29] Z. Zhang, X. Wang, K. Long, A. Vasilakos, and L. Hanzo, "Large-scale MIMO-based wireless backhaul in 5G networks," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 58–66, Oct. 2015.

[30] H. Tabassum, A. H. Sakr, and E. Hossain, "Analysis of massive MIMO-enabled downlink wireless backhauling for full-duplex small cells," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2354–2369, Jun. 2016.

[31] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, and P. Levis, "Applications of self-interference cancellation in 5G and beyond," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 114–121, Feb. 2014.

[32] G. Chen, Y. Gong, P. Xiao, and R. Tafazolli, "Dual antenna selection in self-backhauling multiple small cell networks," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1611–1614, Aug. 2016.

[33] D. Korpi, T. Riihonen, and M. Valkama, "Self-backhauling full-duplex access node with massive antenna arrays: Power allocation and achievable sum-rate," in *Proc. Europ. Signal Process. Conf.*, Aug. 2016, pp. 1618–1622.

[34] "Ten key rules of 5G deployment," Nokia Solutions and Networks, white paper C401-01178-WP-201503-1-EN, 2015.

[35] P. Kela, J. Turkka, and M. Costa, "Borderless mobility in 5G outdoor ultra-dense networks," *IEEE Access*, vol. 3, pp. 1462–1476, Aug. 2015.

[36] X. Huang, K. Yang, F. Wu, and S. Leng, "Power control for full-duplex relay-enhanced cellular networks with QoS guarantees," *IEEE Access*, vol. 5, pp. 4859–4869, Mar. 2017.
[37] A. Sharma, R. K. Ganti, and J. K. Milleth, "Joint backhaul-access analysis of full duplex self-backhauling heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1727–1740, Mar. 2017.
[38] H. Q. Ngo, H. Suraweera, M. Matthaiou, and E. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.
[39] A. Sahai, S. Diggavi, and A. Sabharwal, "On degrees-of-freedom of full-duplex uplink/downlink channel," in *Proc. IEEE Inform. Theory Workshop*, Sep. 2013, pp. 1–5.
[40] K. Kim, S. W. Jeon, and D. K. Kim, "The feasibility of interference alignment for full-duplex MIMO cellular networks," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1500–1503, Sep. 2015.
[41] M. Duarte, A. Feki, and S. Valentin, "Inter-user interference coordination in full-duplex systems based on geographical context information," in *Proc. IEEE Int. Conf. Commun.*, May 2016, pp. 1–7.
[42] H. Yang and T. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
[43] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Nov. 2012, version 20121115. [Online]. Available: http://www2.imm.dtu.dk/pubdb/p.php?3274
[44] I. Rodriguez et al., "Path loss validation for urban micro cell scenarios at 3.5 GHz compared to 1.9 GHz," in *Proc. IEEE Global Commun. Conf.*, Dec. 2013, pp. 3942–3947.

Dani Korpi (S'14) received his B.Sc. and M.Sc. degrees (Hons.) in communications engineering from Tampere University of Technology, Finland, in 2012 and 2014, respectively. He is currently a researcher in the Laboratory of Electronics and Communications Engineering at the same university, pursuing his D.Sc. (Tech.) degree in communications engineering. His main research interest is the study and development of inband full-duplex radios, with a focus on analyzing the RF impairments.

Taneli Riihonen (S'06–M'14) received the D.Sc. degree in electrical engineering (with distinction) from Aalto University, Helsinki, Finland, in August 2014. He is currently an Assistant Professor at the Laboratory of Electronics and Communications Engineering, Tampere University of Technology, Finland. He held various research positions at Aalto University School of Electrical Engineering from September 2005 through December 2017. He was a Visiting Associate Research Scientist and an Adjunct Assistant Professor at Columbia University in the City of New York, USA, from November 2014 through December 2015. He has been nominated eleven times as an Exemplary/Top Reviewer of various IEEE journals and is serving as an Editor for IEEE COMMUNICATIONS LETTERS since October 2014 and for IEEE WIRELESS COMMUNICATIONS LETTERS since May 2017. He received the Finnish technical sector's award for the best doctoral dissertation of the year in Finland within all engineering sciences and the EURASIP Best PhD Thesis Award 2017. His research activity is focused on physical-layer OFDM(A), multiantenna, relaying and full-duplex wireless techniques with current interest in the evolution of beyond 5G systems.

Ashutosh Sabharwal (S'91–M'99–SM'06–F'14) received the B.Tech. degree from Indian Institute of Technology (IIT) Delhi, New Delhi, India, in 1993, and the M.S. and Ph.D. degrees from The Ohio State University, Columbus, OH, USA, in 1995 and 1999, respectively. He is currently a Professor with the Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA. His research interests include information theory, communication algorithms, and the experiment-driven design of wireless networks. He was a recipient of the 1998 Presidential Dissertation Fellowship Award and 2017 Jack Neubauer Memorial Award.

Mikko Valkama (S'00–M'02–SM'15) was born in Pirkkala, Finland, in 1975. He received the M.Sc. and Ph.D. degrees (Hons.) in electrical engineering from the Tampere University of Technology (TUT), Finland, in 2000 and 2001, respectively. In 2002, he received the Best Ph.D. Thesis Award from the Finnish Academy of Science and Letters for his dissertation entitled Advanced I/Q Signal Processing for Wideband Receivers: Models and Algorithms. In 2003, he was a Visiting Post-Doc Researcher with the Communications Systems and Signal Processing Institute, SDSU, San Diego, CA, USA. He is currently a Full Professor and the Laboratory Head with the Laboratory of Electronics and Communications Engineering, TUT. His general research interests include communications signal processing, estimation and detection techniques, signal processing algorithms for flexible radios, cognitive radio, full-duplex radio, radio localization, 5G mobile cellular radio networks, digital transmission techniques such as different variants of multicarrier modulation methods and OFDM, and radio resource management for ad-hoc and mobile networks.

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-4052-3
ISSN 1459-2045